

TITLE OF THE INVENTION: Secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 calcium channel subunit polypeptides and screening assays using same

5 **BACKGROUND OF THE INVENTION**

Voltage-dependent Ca^{2+} channels (VDCCs) are heteromultimeric complexes present in both neuronal and non-neuronal tissues, including heart and skeletal muscle. VDCCs are minimally composed of three subunits: a pore-forming transmembrane α_1 subunit, a hydrophilic intracellular β subunit, and a membrane-associated $\alpha_2\delta$ subunit; a transmembrane γ subunit is also found in skeletal muscle tissue. Multiple subtypes and/or splice variants of the α_1 , β , and $\alpha_2\delta$ subunits have been found.

Gabapentin ((1-aminomethyl)cyclohexane acetic acid or Neurontin) is a structural analogue of GABA, which is mainly used as an adjunctive therapy for epilepsy. Recent research suggests that gabapentin may also have clinical utility for various indications including anxiety and pain. Although designed as a lipophilic GABA-mimetic, gabapentin does not have a high affinity for either GABA_A or GABA_B receptors, GABA uptake sites, or the GABA-degrading enzyme GABA-transaminase (EC 2.6.1.19).

A novel high affinity binding site for [^3H]gabapentin in rat, mouse, and porcine brains has been characterized. Recently, the [^3H]gabapentin-binding protein was isolated from pig brain and identified as the $\alpha_2\delta$ -1 subunit of VDCCs. None of the prototypic anticonvulsant drugs displace [^3H]gabapentin binding from the $\alpha_2\delta$ -1 subunit. [^3H]Gabapentin-binding is stereospecifically inhibited by two enantiomers of 3-isobutyl GABA. The rank order of potency of gabapentin, and S- and R-isobutyl GABA, at the [^3H]gabapentin binding site mirrors their anticonvulsant activity in mice. However, electrophysiological studies have yielded conflicting data on the action of gabapentin at VDCCs.

The $\alpha_2\delta$ subunit is derived from a single gene, the product of which is extensively post-translationally modified particularly through the cleavage of the signal sequence. The polypeptide is cleaved to form disulfide-bridged α_2 and δ peptides, both of which are heavily glycosylated. Although it seems clear today that the α_2 and δ peptides are membrane-associated peptides, it is unclear whether these peptides comprise one or several transmembrane domains. Furthermore, the location, size and structural configuration of these eventual transmembrane domains remains to be determined.

But in any event, the fact that $\alpha_2\delta$ is a membrane-associated protein, regardless of its precise structural configuration, renders its large scale expression in recombinant systems difficult. Indeed, as the $\alpha_2\delta$ protein is targeted to the membrane, it requires detergent solubilisation to release it for purification. This important drawback imposes considerable

5 restrictions for any potential applications requiring large amounts of recombinant protein. Furthermore, the various subtypes of $\alpha_2\delta$ subunits are different proteins with very low homologies. It is therefore extremely difficult to predict their respective behaviors, for example in gene truncation experiments.

10 The only assay currently available for the screening of ligands that bind the $\alpha_2\delta$ subunit involves the use of pig membrane extracts as a source of the $\alpha_2\delta$ subunit. Such an assay presents major inconvenients. Firstly, because the assay material is a membrane extract, it is very difficult to accurately determine the protein composition from one assay preparation to another particularly with regard to the subtype. Also, the presence of various impurities

15 in the assay preparation is a problem in small plate assays. Furthermore, as the protein preparation lacks homogeneity, the interaction between the targeted protein and the assay plate is often quite uneven. This renders the streamlining of the assay in a high throughput format almost impossible to achieve.

20

SUMMARY OF THE INVENTION

In the context of the present invention, the inventors have found that it was possible to delete a portion of the nucleotide sequence encoding a eukaryotic, preferably a mammal

25 cerebral cortical voltage-dependent calcium channel α_2 subunit to yield a soluble secreted protein which retains its affinity for [^3H]gabapentin.

The invention concerns:

- 30 1) A purified or isolated nucleic acid encoding a mammalian secreted soluble cerebral cortical voltage-dependent calcium channel $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide.
- 2) A purified or isolated nucleic acid according to 1), comprising a polynucleotide having at least 90% identity with the sequence encoding :

- Sub D1*
- from amino-acid 1 to between amino-acids 1027 and 1062 of SEQ ID N°20 for $\alpha_2\delta$ -2,
 - from amino-acid 1 to between amino-acids 984 and 1019 of SEQ ID N°22 for $\alpha_2\delta$ -3.

3) A purified or isolated nucleic acid according to 1), having at least 90% identity with the
5 sequence encoding :

- Sub D2*
- from amino-acid 1 to between amino-acids 1047 and 1062 of SEQ ID N°20 for $\alpha_2\delta$ -2,
 - from amino-acid 1 to between amino-acids 1004 and 1019 of SEQ ID N°22 for $\alpha_2\delta$ -3.

Sub D3

4) A purified or isolated nucleotide sequence according to 1) wherein said sequence is the
sequence of SEQ ID N°1, SEQ ID N°2, SEQ ID N°3, SEQ ID N°7, SEQ ID N°8, SEQ ID
N°9, SEQ ID N°13, SEQ ID N°14, SEQ ID N°15, SEQ ID N°19 or SEQ ID N°21.

Sub D4

5) A purified or isolated nucleic acid, having at least 90% identity with the nucleotide
sequence of SEQ ID N°19 or SEQ ID N°21.

15

Sub D5

6) A purified or isolated polynucleotide comprising at least 10 consecutive nucleotides of
the nucleotide sequence of SEQ ID N°19 or SEQ ID N°21.

Sub D6

7) A polynucleotide probe or primer hybridizing, under stringent conditions, with the with
the nucleotide sequence of SEQ ID N°19 or SEQ ID N°21.

8) A method for the amplification of a nucleic acid encoding a mammalian secreted
soluble cerebral cortical voltage-dependent calcium channel $\alpha_2\delta$ -n subunit polypeptide
wherein n is 2, 3 or 4, said method comprising the steps of:

- 25
- (a) contacting a test sample suspected of containing the target secreted soluble $\alpha_2\delta$ -
n subunit nucleic acid, or a sequence complementary thereto, with an amplification
reaction reagent comprising a pair of amplification primers located on either side of
the $\alpha_2\delta$ -n subunit nucleic acid region to be amplified, and
 - (b) optionally, detecting the amplification products.

30

9) A kit for the amplification of a nucleic acid encoding a secreted soluble $\alpha_2\delta$ -n subunit
polypeptide wherein n is 2, 3 or 4, or a complementary sequence thereto in a test sample,
wherein said kit comprises:

(a) a pair of oligonucleotide primers which can hybridize, under stringent conditions, to the secreted soluble $\alpha_2\delta$ -n subunit nucleic acid region to be amplified;

(b) optionally, the reagents necessary for performing the amplification reaction.

5

10) A recombinant vector comprising a nucleic acid according to 1).

11) A recombinant host cell comprising a nucleic acid according to 1).

10 12) A method for producing a secreted soluble $\alpha_2\delta$ -n subunit wherein n is 2, 3 or 4, and said method comprises the steps of:

(a) inserting the nucleic acid encoding the desired $\alpha_2\delta$ -n subunit polypeptide in an appropriate vector;

15

(b) culturing, in an appropriate culture medium, a host cell previously transformed or transfected with the recombinant vector of step (a);

(c) harvesting the culture medium thus obtained or lyse the host cell, for example by sonication or osmotic shock;

(d) separating or purifying, from said culture medium, or from the pellet of the resultant host cell lysate, the thus produced $\alpha_2\delta$ -n subunit polypeptide of interest.

20

13) A purified or isolated recombinant polypeptide comprising the amino acid sequence of a secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide.

25

14) A recombinant polypeptide according to 14), having at least 80% amino-acid identity with a polypeptide comprising :

- from amino acid 1 to between amino acids 1027 and 1062 of the amino acid sequence of SEQ ID N°20, or

- from amino acid 1 to between amino acids 1019 and 1079 of the amino acid sequence of SEQ ID N°22.

30

15) A recombinant polypeptide according to 14), wherein said recombinant polypeptide is selected from the group consisting of the amino acid sequences of SEQ ID n°4, SEQ ID n°5, SEQ ID n°6, SEQ ID n°10, SEQ ID n°11, SEQ ID n°12, SEQ ID n°16, SEQ ID n°17 and SEQ ID n°18.

16) A method for the screening of ligands which bind a cerebral cortical voltage-dependent calcium channel $\alpha_2\delta$ -n subunit wherein n is 2, 3 or 4, said method comprising the steps of:

- 5 - contacting a secreted soluble recombinant calcium channel $\alpha_2\delta$ -n subunit polypeptide with:
- a ligand of interest; and
 - a labelled compound which binds the $\alpha_2\delta$ -n subunit; and
 - measuring the level of binding of the labelled compound to the $\alpha_2\delta$ -n subunit.

10

17) A method according to 16), wherein said method is a scintillation proximity assay.

18) A method according to 16), wherein said method is a flashplate assay.

15 19) A method according to 16), wherein said method is a filter binding assay.

20) A method according to 16), wherein said secreted soluble recombinant calcium channel $\alpha_2\delta$ -n subunit polypeptide is selected from polypeptides having at least 80%, preferably 90%, more preferably 95%, and most preferably 98 or 99% amino-acid identity with the polypeptide comprising from amino acid 1 to between amino-acids 984 and 1063, preferably between amino-acids 994 and 1054, and most preferably between amino-acids 1019 and 1054 of SEQ ID N°5 or SEQ ID N°16.

21) A method according to 16), wherein said secreted soluble recombinant calcium channel $\alpha_2\delta$ -n subunit polypeptide is selected from the group consisting of SEQ ID N°6, 7, 8, 9, 13, 14 and 15, with the polypeptides of SEQ ID N°9 and SEQ ID N°15 being most preferred.

22) A kit for the screening of ligands which bind a cerebral cortical voltage-dependent calcium channel $\alpha_2\delta$ -n subunit wherein n is 2, 3 or 4, said kit comprising:

- 30
- a secreted soluble recombinant calcium channel $\alpha_2\delta$ -n subunit; and
 - a labelled compound which binds to the $\alpha_2\delta$ -n subunit.

Hence, the invention concerns nucleotide sequence fragments of a cerebral cortical voltage dependent calcium channel $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit cDNA encoding a soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide (hereinafter a $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit). Preferably, these nucleotide sequences encode a soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide bearing a gabapentin or a [3 H]gabapentin binding site. More preferably, the soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit nucleic acid is derived from a eukaryotic, preferably a mammal, more preferably a human $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit.

10

A further object of the present invention concerns recombinant vectors comprising a nucleic acid sequence encoding a soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide.

The invention also encompasses host cells and transgenic non-human mammals comprising said nucleic acid sequences or recombinant vectors.

15

The invention also concerns a soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide which is characterized in that it is a soluble secreted polypeptide having affinity for [3 H]gabapentin. Preferably, the soluble secreted polypeptide is derived from a mammal, more preferably a human $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit.

20

The inventors have also found that it was possible to use a soluble secreted form of a voltage-dependant calcium channel $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide in an assay for the screening of ligands which bind the $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit.

25

The invention therefore also concerns a method for the screening of ligands which bind a calcium channel $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit.

The method comprises the steps of:

30

- contacting a secreted soluble recombinant calcium channel $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide with:
 - a ligand of interest; and
 - a labelled compound which binds a $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit; and
- measuring the level of binding of the labelled compound to the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit.

35

The invention also concerns a kit for the screening of ligands which bind a calcium channel $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit.

The kit comprises:

- a secreted soluble recombinant calcium channel $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide; and
- a labelled compound which binds a calcium channel $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit.

10 DETAILED DESCRIPTION OF THE INVENTION

The invention concerns truncated $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit cDNA sequences. These truncated sequences encode soluble secreted polypeptides which retain their affinity for [3 H]gabapentin.

Throughout the present specification, the expression "nucleotide sequence" is used to designate indifferently a polynucleotide or a nucleic acid. More precisely, the expression "nucleotide sequence" encompasses the nucleic material and the sequence information and is not restricted to the sequence information (i.e. the succession of letters chosen among the four base letters) that biochemically characterizes a specific DNA or RNA molecule.

As used interchangeably herein, the terms "oligonucleotides", "nucleic acids" and "polynucleotides" include RNA, DNA, or RNA/DNA hybrid sequences of more than one nucleotide in either single chain or duplex form.

Further to its general meaning understood by the one skilled in the art, the term "nucleotide" is also used herein to encompass modified nucleotides which comprise at least one of the following modifications (a) an alternative linking group, (b) an analogous form of purine, (c) an analogous form of pyrimidine, or (d) an analogous sugar. For examples of analogous linking groups, purines, pyrimidines, and sugars, see for example PCT publication N^oWO 95/04064.

The polynucleotide sequences of the invention may be prepared by any know method, including synthetic, recombinant, or a combination thereof as well as through any purification methods known in the art.

A) Secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides

The invention comprises polynucleotide sequences encoding a soluble secreted eukaryotic, preferably a soluble secreted mammal $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide. These

sequences particularly include but are not restricted to 1) those sequences encoding a soluble secreted polypeptide of this $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit which preferably retains its binding affinity for [^3H]gabapentin and 2) nucleotide fragments useful as nucleic acid primers or probes for amplification or detection purposes.

- 5 The expression "soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit" is intended to designate polypeptide sequences which, when produced by a recombinant host cell, are secreted at least partially into the culture medium rather than remaining associated with the host cell membrane.

10 **1) cDNA fragments encoding soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3, $\alpha_2\delta$ -4 subunit polypeptides**

One of the important embodiments of the present invention concerns truncated nucleotide sequences of $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit cDNAs which encode soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides. The inventors have found that it was possible to
 15 generate deletion mutants of $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit cDNAs which, when expressed, produce a significant amount of soluble secreted proteins, preferably soluble secreted proteins, which retain their [^3H]gabapentin binding affinity. These truncated nucleotide sequences of the invention are of significant value to the skilled person as they now allow fast and reliable access to significant concentrations of selected soluble secreted $\alpha_2\delta$ -2,
 20 $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides. To that end, the inventors have determined the minimal and optimal fragment lengths required to express a polypeptide which: 1) binds [^3H]gabapentin with sufficient affinity and; 2) is obtained in a soluble secreted form.

The discussion provided below provides comments on possible truncations, giving as an example the human $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit. However, given the very substantial
 25 cross-species homology for $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit sequences, the comments below can also be applied to other eukaryotic species, and more particularly other mammalian species such as rat, mouse, rabbit or pig. Their $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit sequences, which for most are available in public databases, share a very substantial homology with the human $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit sequences.

30

The inventors believe ~~that~~ the soluble secreted $\alpha_2\delta$ -2 subunit polypeptides which are as close as possible to the native sequence and which are therefore more likely to retain their native folding and hence their [^3H]gabapentin binding properties are those corresponding to the native protein in which amino-acid stretch 1027 to the C-terminal end of
 35 the amino-acid sequence of SEQ ID N°20 has been deleted. The skilled scientist can quite easily determine within this amino-acid stretch the optimal $\alpha_2\delta$ -2 subunit polypeptides.

The inventors also believe that the soluble secreted $\alpha_2\delta$ -3 subunit polypeptides which are as close as possible to the native sequence and which are therefore more likely to retain

their native folding and hence their $[^3\text{H}]$ gabapentin binding properties are those corresponding to the native protein in which amino-acid stretch 984 to C-terminal end of the amino-acid sequence of SEQ ID N°22 has been deleted. The skilled scientist can quite easily determine within this amino-acid stretch the optimal $\alpha_2\delta$ -3 subunit polypeptides.

The invention therefore particularly concerns a nucleotide sequence encoding a polypeptide having at least 80% identity with the polypeptide comprising from amino-acid 1 to between amino-acids 1027 and 1145, preferably to between amino-acids 1062 and 1145 of SEQ ID N°20.

Preferred nucleotide sequences include those of SEQ ID N°1, SEQ ID N° 2 and SEQ ID N°3.

The invention also concerns a nucleotide sequence encoding a polypeptide having at least 80% identity with the polypeptide comprising from amino-acid 1 to between amino-acids 984 and 1085, preferably to between amino-acids 1019 and 1085 of SEQ ID N°22.

Preferred nucleotide sequences include those of SEQ ID N°7, SEQ ID N° 8 and SEQ ID N°9.

B) Amplification of the soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit nucleotide sequences

Another object of the invention consists of a method for the amplification of a nucleic acid encoding a soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide, preferably a polypeptide bearing a $[^3\text{H}]$ gabapentin binding site, said method comprising the steps of:

(a) contacting a test sample suspected of containing the target $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit nucleic acid, a fragment or a variant thereof, or a sequence complementary thereto, with an amplification reaction reagent comprising a pair of amplification primers which can hybridize under stringent conditions, the $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit nucleic acid region to be amplified, and

(b) optionally, detecting the amplification products.

The expression $[^3\text{H}]$ gabapentin binding site, when used herein is intended to designate a site which can bind either $[^3\text{H}]$ gabapentin or other ligands such as (S+)-3-isobutyl gaba or (R-)-3-isobutyl gaba.

In a first preferred embodiment of the above method, the nucleic acid encodes a secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide of SEQ ID N°4, SEQ ID N°5, SEQ ID N°6, SEQ ID N°10, SEQ ID N°11, SEQ ID N°12, SEQ ID N°16, SEQ ID N°17 and SEQ ID N°18.

In a second preferred embodiment of the above amplification method, the amplification product is detected by hybridization with a labelled probe having a sequence which is complementary to the amplified region.

5

C) Recombinant vectors and hosts cells for the expression of a secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide

1) Recombinant vectors

10 The present invention also encompasses a family of recombinant vectors comprising any one of the nucleic acids described herein. Firstly, the invention deals with a recombinant vector comprising a nucleic acid selected from the group consisting of:

(a) a purified or isolated nucleic acid encoding a a secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit having at least 80% amino acid identity with the polypeptide of SEQ ID N°20 or 22, or a sequence complementary thereto;

(b) a purified or isolated nucleic acid having at least 90% nucleotide identity with a polynucleotide selected from the group consisting of the nucleotide sequences of SEQ ID N°1, SEQ ID N°2, SEQ ID N°3, SEQ ID N°7, SEQ ID N°8, SEQ ID N°9, SEQ ID N°13, SEQ ID N°14, SEQ ID N°15 or a sequence complementary thereto;

(c) a purified or isolated polynucleotide comprising at least 10 consecutive nucleotides of a nucleic acid described in (a) or (b) or a sequence complementary thereto.

In a first preferred embodiment a recombinant vector of the invention is used to amplify the inserted polynucleotide of the invention in a suitable host cell, this polynucleotide being amplified every time the recombinant vector replicates.

Recombinant expression vectors comprising a nucleic acid encoding secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides that are described in the present specification are also part of the invention. These include, but are not restricted to, nucleic acids encoding from amino-acid 1 to between amino-acids 1027 and 1145, preferably between amino-acids 1062 and 1145 of SEQ ID N°20, as well as nucleic acids encoding from amino-acid 1 to between amino-acids 984 and 1085, preferably between amino-acids 1019 and 1085, of SEQ ID N°22.

Another preferred embodiment of the recombinant vectors according to the invention consist of expression vectors comprising a nucleic acid encoding $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides of the invention, and more preferably a nucleic acid encoding a polypeptide selected from the group consisting of the amino acid sequences of SEQ ID N°4, SEQ ID N°5, SEQ ID N°6, SEQ ID N°10, SEQ ID N°11, SEQ ID N°12, SEQ ID N°16, SEQ ID N°17 and SEQ ID N°18.

Within certain embodiments, expression vectors can be employed to express the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides which can then be purified and for example, be used as an immunogen in order to raise specific antibodies directed against said secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides.

- 5 Preferred eukaryotic vectors of the invention are listed hereafter as illustrative but not limitative examples: pcDNA3, pFLAG, pCMV-Script, pIND, pMC1NEO, pHIL, pGAPZA, pMT/V5-His-TOPO, pMT/V5-His, pAc5.1/V5-HisA, pDS47/V5-His, pcDNA4, pcDNA6, pEF1, pEF4, pEF6, pUB6, pZeoSV2, pRc/CMV2, pcDM8, pCR3.1, pDisplay, pSecTag2, pVP22, pEMZ, pCMV/Zeo, pSinRep5, pCEP, pREP, pHook-1

10

Preferred bacteriophage recombinant vectors of the invention are P1 bacteriophage vectors such as described by Sternberg N.L. (1992;1994).

- A suitable vector for the expression of a soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide is a baculovirus vector that can be propagated in insect cells and in insect cell-lines. Specific suitable host vectors includes, but are not restricted to : pFastBac-1, pIZ/V5-His, pBacMan-1, pBlueBac4.5, pBlueBacHis2, pMelBacA, pVL1392, pVL1393

15

The recombinant expression vectors from the invention may also be derived from an adenovirus such as those described by Feldman and Steig. (1996) or Ohno et al. (1994).

20

Another preferred recombinant adenovirus according to this specific embodiment of the present invention is the human adenovirus type two or five (Ad 2 or Ad 5) or an adenovirus of animal origin (French Patent Application n°FR 93 05 954).

a) Regulatory expression sequences

25

Expression requires that appropriate signals are provided in the vectors, said signals including various regulatory elements such as enhancers/promoters from both viral and mammalian sources that drive expression of the genes of interest in host cells. The regulatory sequences of the expression vectors of the invention are operably linked to the nucleic acid a soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide.

30

As used herein, the term "operably linked" refers to a linkage of polynucleotide elements in a functional relationship. For instance, a promoter or an enhancer is operably linked to a coding sequence if it affects the transcription of the coding sequence.

- More precisely, two DNA molecules (such as a polynucleotide containing a promoter region and a polynucleotide encoding a desired polypeptide or polynucleotide) are said to be "operably linked" if the nature of the linkage between the two polynucleotides does not : (1) result in the introduction of a frame-shift mutation or (2) interfere with the ability of the polynucleotide containing the promoter to direct the transcription of the coding polynucleotide.

35

Generally, recombinant expression vectors include origins of replication, selectable markers permitting transformation of the host cell, and a promoter derived from a highly

40

expressed gene to direct transcription of a downstream structural sequence. The heterologous structural sequence is assembled in an appropriate frame with the translation, initiation and termination sequences, and preferably a leader sequence capable of directing sequences of the translated protein into the periplasmic space or the extra-cellular medium.

- 5 In a specific embodiment wherein the vector is adapted for transfecting and expressing desired sequences in eukaryotic host cells, preferred vectors comprise an origin of replication from the desired host, a suitable promoter and an enhancer, and also any necessary ribosome binding sites, polyadenylation site, transcriptional termination sequences, and optionally 5'-flanking non-transcribed sequences.
- 10 DNA sequences derived from the SV 40 viral genome, for example SV 40 origin early promoter, enhancer, and polyadenylation sites may be used to provide the required non-transcribed genetic elements.

b) Promoter sequences

- 15 Suitable promoter regions used in the expression vectors according to the invention are chosen taking into account the host cell in which the heterologous nucleic acids have to be expressed.

A suitable promoter may be heterologous with respect to the nucleic acid for which it controls the expression, or alternatively can be endogenous to the native polynucleotide containing the coding sequence to be expressed.

- 20 Additionally, the promoter is generally heterologous with respect to the recombinant vector sequences within which the construct promoter/coding sequence has been inserted. Preferred eukaryotic promoters are the CMV, polyhidran or OPIE2.

c) Recombinant host cells

- 25 Host cells that have been transformed or transfected with one of the nucleic acids described herein, or with one of the recombinant vector, particularly recombinant expression vector, described herein are also part of the present invention.

- Are included host cells that are transformed (prokaryotic cells) or are transfected (eukaryotic cells) with a recombinant vector such as one of those described above. Preferred host cells used as recipients for the expression vectors of the invention are the following:

- (a) prokaryotic host cells: *Escherichia coli*, strains. (i.e. DH10 Bac strain) *Bacillus subtilis*, *Salmonella typhimurium* and strains from species such as *Pseudomonas*, *Streptomyces* and *Staphylococcus*;

- (b) eukaryotic host cells: HeLa cells (ATCC N°CCL2; N°CCL2.1; N°CCL2.2), Cv 1 cells (ATCC N°CCL70), COS cells (ATCC N°CRL 1650; N°CRL 1651), Sf-9 cells (ATCC N°CRL 1711), C127 cells (ATCC N°CRL-1804), 3T3 cells (ATCC N°CRL-6361),

CHO cells (ATCC N°CCL-61), human kidney 293 cells (ATCC N° 45504; N°CRL-1573), BHK (ECACC N°84100 501; N°84111301), sf 9, sf 21 and hi-5 cells.

D) Production of recombinant secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides

5 The present invention also concerns a method for producing one of the amino acid sequences described herein and especially a polypeptide selected from the group consisting of the amino acid sequences of SEQ ID N°4, SEQ ID N°5, SEQ ID N°6, SEQ ID N°10, 10 SEQ ID n°11, SEQ ID n°12, SEQ ID n°16, SEQ ID n°17 or SEQ ID n°18 wherein said method comprises the steps of:

(a) inserting the nucleic acid encoding the desired amino acid sequence in an appropriate vector;

15 (b) culturing, in an appropriate culture medium, a host cell previously transformed or transfected with the recombinant vector of step (a);

(c) harvesting the culture medium thus obtained or lyse the host cell, for example by sonication or osmotic shock;

20 (d) separating or purifying, from said culture medium, or from the pellet of the resultant host cell lysate, the thus produced recombinant polypeptide of interest.

In some instances, it is required to tag the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide prior to purification. The tag is then in most instances encoded into the nucleotide sequence that is needed to express the polypeptide. Examples of such tags include, but are not limited to sequences encoding C-myc, FLAG, a sequence of histidine residues, 25 heamagglutinin A, V5, Xpress or GST. Most of these tags can be incorporated directly into the sequence, for instance through PCR amplification by incorporating the appropriate coding sequence in one of the PCR amplification primers. However, the tag can also be introduced by other means such as covalent binding of the appropriate nucleic acid sequence encoding the tag moiety with the 3' or 5' end of the nucleic acid sequence 30 encoding the polypeptide sequence. This is the case for GST.

Purification of the recombinant secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3, $\alpha_2\delta$ -4 subunit polypeptides according to the present invention is then carried out by passage onto a nickel or copper affinity chromatography column, such as a Ni NTA column or a Q-Sepharose column.

In another embodiment of the above method, the polypeptide thus produced is further characterized, for example by binding onto an immuno-affinity chromatography column on which polyclonal or monoclonal antibodies directed to the secreted soluble $\alpha_2\delta$ -2 subunit polypeptide of interest have been previously immobilised.

5

In another embodiment of the invention, the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3, $\alpha_2\delta$ -4 subunit polypeptide can be only partially purified. For instance, it can be purified along with other contaminating proteins using an appropriate chromatography matrix such as an ion-exchange chromatography column. In such instances, it is not required to tag the desired

10

The most preferred embodiment contemplated by the inventors concerns the use of a purified tagged secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide. A particularly preferred tag is a nucleotide sequence encoding from 2 to 10, and preferably 6 histidine

15

residues.

With regard to the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide used subsequently in the screening assay of the invention, several possibilities are also open to the skilled person.

20

In a first and preferred embodiment, the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide comprises a tag moiety which can be selected among the tags referred to above. Such tagged polypeptides are particularly useful in SPA or flashplate assays. A preferred tag is the nucleotide sequence encoding histidine residues referred to above.

25

In a second embodiment, the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide can be used without a tag. This is the case for instance in SPA or flashplate assays comprising beads or plates coated with wheat germ lectin. In such an embodiment, the tag is not needed as the carbohydrate moieties of the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4

30

subunit polypeptide bind directly to the wheat germ lectin-coated beads or plates.

1) Labelled compounds which bind the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide

In cases where the $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 binding site is the [^3H]gabapentin binding site, the preferred labelled compound which can be used is of course gabapentin itself.

- 5 However, gabapentin is not the only labelled compound which can be used in this context. Indeed, it has been previously demonstrated that saturation binding analyses on porcine synaptic plasma cerebral cortex membranes performed in the presence of L-leucine indicate a competitive interaction of the amino acid with the [^3H]gabapentin binding site, significantly reducing [^3H]gabapentin binding affinity for the site. The inventors believe
- 10 that this competitive interaction is true across across all the amino-acids listed in table 1 below.

15 **TABLE 1**

Binding affinities of selected amino acids ($\text{IC}_{50} < 500\text{nM}$) for the [^3H]gabapentin site in porcine cortical membranes

	<u>COMPOUND</u>	<u>IC_{50} (NM) ARITHMETIC MEAN (N=3) \pm S.E.M.</u>
20	Gabapentin	42.1 ± 5.5
	L-Norleucine	23.6 ± 6.7
	L-Allo-Isoleucine	32.8 ± 6.0
	L-Methionine	49.6 ± 10.0
	L-Leucine	61.3 ± 20.9
25	L-Isoleucine	68.8 ± 1.9
	L-Valine	330 ± 18
	L-Phenylalanine	351 ± 89

- It is therefore possible to use commercially available labelled forms of these high affinity
- 30 ligands in replacement of gabapentin. The utility of [^3H]L-leucine has been demonstrated in a filter binding assay and in a flashplate assay format. The inventors believe that labelled amino acids but also other compounds, with affinities preferably below 500 nM in the binding assay can be used as replacements of gabapentin.

With regard to the label, several embodiments can be used in the context of the invention. Preferred labels are of course radioactive labels, a list of which is provided further in this specification.

5

2) Assay formats and conditions

Several assay formats can be used to carry out the method of the present invention. Preferred assay formats include scintillation assays such as the scintillation proximity assay (SPA) or the flashplate assay. Other assay formats well known to those skilled in the arts such as the filter binding assay and the centrifugation assay are also contemplated in the present invention.

SPA and flashplate assays are preferred assay formats for the present invention. Additional details on these assays are provided below.

15 Scintillation assay format

Scintillation assays technology either involves the use of scintillant beads (for the SPA assay) or plates (for the flashplate assay). SPA beads are usually made from either cerium-doped yttrium ion silicate ($\text{Y}_2\text{SiO}_5:\text{Ce}$) or polyvinyltoluene (PVT) containing an organic scintillant such as PPO. Flashplates commonly used are those such as Ni chelate flashplates although other flashplates can also be used.

Assays are usually carried out in aqueous buffers using radioisotopes such as ^3H , ^{125}I , ^{14}C , ^{35}S or ^{33}P that emit low-energy radiation, the energy of which is easily dissipated in an aqueous environment. For example, the electrons emitted by ^3H have an average energy of only 6 keV and have a very short path length ($\sim 1 \mu\text{m}$) in water. If a molecule labelled with one of these isotopes is bound to the bead or flashplate surface, either directly or via interaction with another molecule previously coupled to the bead or flashplate, the emitted radiation will activate the scintillant and produce light. The amount of light produced, which is proportional to the amount of labelled molecules bound to the beads, can be measured conveniently with a liquid scintillation (LS) counter. If the labelled molecule is not attached to the bead or a flashplate surface, its radiation energy is absorbed by the surrounding aqueous solvent before it reaches the bead, and no light is produced. Thus, bound ligands give a scintillation signal, but free ligands do not, and the need for a time-consuming separation step, characteristic of conventional radioligand binding assays, is

09397550-091609

The conditions under which SPA and flashplate assays are performed in the context of the present invention are provided below.

Scintillation assay conditions

The SPA assays is first developed to optimize the conditions under which the radioligand binds the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide. The parameters which can be varied to optimize radioligand binding in a typical SPA assay using Amersham beads include assay temperature, $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide interaction with the radioligand and the SPA beads, radioligand concentration as well as pH variations.

15 The temperature at which the assay can be carried out can vary from 1 to 30°C. Preferred temperatures range from 18 to 23°C, with 21°C being the most preferred temperature. The interaction of the $\alpha_2\delta$ subunit polypeptide with the SPA beads can be optimized by adjusting the concentration of the polypeptide and by introducing a reagent which will favor this interaction. When 50 mg of Amersham SPA beads are used, the $\alpha_2\delta$ -1 subunit polypeptide concentration may vary from 0.1 to 10 pmoles per well, with the optimal concentration being generally around 5 to 6 pmoles per well.

As for the reagent favoring the interaction between the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide and the radioligand as well as the Amersham SPA beads, the
25 inventors found that imidazole could be efficiently used for that purpose when the $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide is tagged with an amino acid sequence including 6 histidine residues. Furthermore, and more importantly, it was found that imidazole also enhanced binding of the radioligand to the $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 polypeptide.

30 The concentration of the radioligand is evaluated with respect to the concentration of secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide present in the assay medium. Generally, the concentration of radioligand varies from 1 nM to 100 nM. A preferred [^3H]gabapentin concentration is about 5 to 20 nM, with a most preferred concentration

being about 10 nM. A preferred [^3H]leucine concentration is also about 5 to 20 nM, with a most preferred concentration being about 10 nM. It is to be noted that the concentration of other radioligands having affinities similar to those of [^3H]gabapentin and [^3H]leucine should also be in the range of about 5 to 20 nM.

5

Once the optimal radioligand binding conditions have been determined, a test ligand can be introduced in the assay medium to evaluate the level of displacement of the radioligand. The concentration of test ligand to be introduced in the assay medium usually varies from 0.1 nM to about 100 μM . A preferred test ligand concentration of about 10 μM is usually a starting point in a high throughput screening assay. Then, depending on the number of hits obtained, it may be lowered or increased.

10

It is to be noted that the parameters set forth above, which have been evaluated for a typical SPA assay using Amersham SPA beads can be adjusted by the skilled person, for example if SPA beads of a different type are used.

15

b) Flashplate assay

Similarly to the SPA assays, the flashplate can first be developed in order to optimize the conditions under which the radioligand binds the $\alpha_2\delta$ subunit polypeptide. The parameters which can be varied to optimize radioligand binding in a typical flashplate assay using NEN Ni chelate flashplates also include assay temperature, secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide interaction with both the radioligand and the flashplates, radioligand concentration as well as pH variations.

20

The temperature at which the assay can be carried out can vary from 1 to 30°C. Preferred temperatures range from 18 to 23°C, with 21°C being the most preferred temperature.

25

The interaction of the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide with the flashplates can be optimized by adjusting the concentration of the polypeptide and by introducing a reagent which will favor this interaction. When a standard NEN Ni chelate flashplate is used, the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide volume usually varies between 0.5 and 20 μl for a concentration of secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide of 0.6 pmol/ μl . As the published maximum binding capacity of

30

2ub
b7 NEN p plates is about 6 pmol per well, the inventors consider that an optimal concentration of secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide is probably around 5 pmol per well at 8 ul.

- 2ub
b7
- 5 With regard to the reagent favoring the interaction between the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide and the radioligand as well as the flashplates, the inventors believe that imidazole could also be efficiently used for that purpose when the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide is tagged with an amino acid sequence including 6 histidine residues. The inventors also believe that imidazole concentrations can substantially enhanced binding of the radioligand to the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 polypeptide. The optimal concentration of imidazole used to enhance radioligand binding varies depending on the concentration of secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide used in the assay. For instance, when the volume of the $\alpha_2\delta$ -1 subunit polypeptide is about 10 ul ul ($\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 polypeptide concentration of 0.6 pmol/ul), the optimal imidazole concentration can vary between 1 and 20 mM, with a concentration of about 10 mM being preferred. As mentioned previously, other compounds such as histidine as well as pH variations may be used to enhance radioligand binding.
- 15
- 20 The concentration of the radioligand is evaluated with respect to the concentration of $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide present in the assay medium. Generally, the concentration of radioligand varies from 1 nM to 100 nM. A preferred [3 H]gabapentin concentration is about 5 to 20 nM, with a most preferred concentration being about 10 nM. A preferred [3 H]leucine concentration is also about 5 to 20 nM, with a most preferred concentration being about 10 nM. It is to be noted that the concentration of other radioligands having affinities similar to those of [3 H]gabapentin and [3 H]leucine should also be in the range of about 5 to 20 nM.
- 25

Once the optimal radioligand binding conditions have been determined, a test ligand can be introduced in the assay medium to evaluate the level of displacement of the radioligand. The concentration of test ligand to be introduced in the assay medium usually varies from 0.1 nM to about 100 uM. A preferred test ligand concentration of about 10 uM is usually a

30

2ub
b7

starting point in a high throughput screening assay. Then, depending on the number of hits obtained, it may be lowered or increased.

The inventors have tested the displacement of a particular radioligand, [³H]gabapentin, with (S+)-3-isobutyl gaba. The data provided in the examples which follow clearly shows that the assay can be used in high throughput competition studies.

E) Purified recombinant secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 polypeptides

Another object of the present invention consists of a purified or isolated recombinant polypeptide comprising the amino acid sequence of a secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide.

Preferred isolated recombinant polypeptides of the invention include those having at least 80%, preferably 90%, more preferably 95, and most preferably 98 or 99%, amino-acid identity with polypeptides comprising from amino acid 1 to between amino-acids 1027 and 1145, preferably between amino-acids 1062 and 1145 of SEQ ID N°20, as well as those having at least 80%, preferably 90%, more preferably 95, and most preferably 98 or 99%, amino-acid identity with polypeptides comprising from amino acid 1 to between amino-acids 984 and 1085, preferably between amino-acids 1019 and 1085 of SEQ ID N°22.

In a further preferred embodiment, the polypeptide comprises an amino acid sequence having at least 80%, preferably 90%, more preferably 95%, and most preferably 98% or 99% amino acid identity with the amino acid sequence of SEQ ID N°4, SEQ ID N°5, SEQ ID N°6, SEQ ID N°10, SEQ ID N°11, SEQ ID N°12, SEQ ID N°16, SEQ ID N°17 and SEQ ID N°19.

F) Modified secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides

The invention also relates to secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide comprising amino acid changes ranging from 1, 2, 3, 4, 5, 10, 20, 25, 30, 35, 40 substitutions, additions or deletions of one amino acid as regards to polypeptides of anyone of the amino acid sequences of the present invention. Preferred sequences are those of SEQ ID N°4, SEQ ID N°5, SEQ ID N°6, SEQ ID N°10, SEQ ID N°11, SEQ ID n°12, SEQ ID n°16, SEQ ID n°17 and SEQ ID N°18.

In the case of an amino acid substitution in the amino acid sequence of a polypeptide according to the invention, one or several consecutive or non-consecutive amino-acids are

replaced by "equivalent" amino-acids. The expression "equivalent" amino acid is used herein to designate any amino acid that may be substituted for one of the amino-acids belonging to the native protein structure without decreasing the binding properties of the corresponding peptides to the antibodies raised against the polypeptides of the invention.

- 5 In other words, the "equivalent" amino-acids are those which allow the generation or the synthesis of a polypeptide with a modified sequence when compared to the amino acid sequence of the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides of interest, said modified polypeptide being able to bind to the antibodies raised against the secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide of interest and/or to induce antibodies
- 10 recognizing the parent polypeptide.

- Alternatively, amino acid changes encompassed are those which will not abolish the biological activity of the resulting modified polypeptide. These equivalent amino-acids may be determined either by their structural homology with the initial amino-acids to be replaced, by the similarity of their net charge or of their hydrophobicity, and optionally by
- 15 the results of the cross-immunogenicity between the parent peptides and their modified counterparts.

The peptides containing one or several "equivalent" amino-acids must retain their specificity and affinity properties to the biological targets of the parent protein, as it can be assessed by a ligand binding assay or an ELISA assay.

- 20 Examples of amino-acids belonging to specific classes include Acidic (Asp, Glu), Basic (Lys, Arg, His), Non-polar (Ala, Val, Leu, Ile, Pro, Met, Phe, Trp) or uncharged Polar (Gly, Ser, Thr, lys, Tyr, Asn, Gln) amino-acids.

- Preferably, a substitution of an amino acid in $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide of the invention, or in a peptide fragment thereof, consists in the replacement of an amino
- 25 acid of a particular class for another amino acid belonging to the same class.

By an equivalent amino acid according to the present invention is also contemplated the replacement of a residue in the L-form by a residue in the D form or the replacement of a Glutamic acid (E) residue by a Pyro-glutamic acid compound. The synthesis of peptides containing at least one residue in the D-form is, for example, described by Koch (1977).

- 30 A specific embodiment of a modified peptide of interest according to the present invention, includes, but is not limited to, a peptide molecule, which is resistant to proteolysis. This is a peptide in which the -CONH- peptide bond is modified and replaced by a (CH₂NH) reduced bond, a (NHCO) retro inverso bond, a (CH₂-O) methylene-oxy bond, a (CH₂S) thiomethylene bond, a (CH₂CH₂) carba bond, a (CO-CH₂) cetomethylene bond, a (CHOH-CH₂) hydroxyethylene bond), a (N-N) bound, a E-alcene bond or also a -CH=CH-bond.
- 35

The invention also encompasses secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide in which at least one peptide bond has been modified as described above.

009750-051500

The polypeptides according to the invention may also be prepared by the conventional methods of chemical synthesis, either in a homogenous solution or in solid phase. As an illustrative embodiment of such chemical polypeptide synthesis techniques, it may be cited the homogenous solution technique described by Houbenweyl (1974).

- 5 The secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide of interest, or a fragment thereof may thus be prepared by chemical synthesis in liquid or solid phase by successive couplings of the different amino acid residues to be incorporated (from the N-terminal end to the C-terminal end in liquid phase, or from the C-terminal end to the N-terminal end in solid phase) wherein the N-terminal ends and the reactive side chains are previously
10 blocked by conventional groups.
- For solid phase synthesis, the technique described by Merrifield (1965a; 1965b) may be used in particular.

G) Antibody production

- 15 The secreted soluble $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptides of the invention and their peptide fragments of interest can be used for the preparation of antibodies.
- Polyclonal antibodies may be prepared by immunization of a mammal, especially a mouse or a rabbit, with a polypeptide according to the invention that is combined with an adjuvant
20 of immunity, and then by purifying the specific antibodies contained in the serum of the immunized animal on an affinity chromatography column on which has previously been immobilized the polypeptide that has been used as the antigen.
- Monoclonal antibodies may be prepared from hybridomas according to the technique described by Kohler and Milstein (1975).
- 25 The present invention also deals with antibodies produced by the trioma technique and by the human B-cell hybridoma technique, such as described by Kozbor et al. (1983).
- Antibodies of the invention also include chimeric single chain Fv antibody fragments (US Patent N° 4,946,778; Martineau et al., (1998), antibody fragments obtained through phage display libraries Ridder et al. (1995) and humanized antibodies (Leger et al., (1997)).

H) Screening assays

- 30 The invention concerns a method for the screening of ligands which bind soluble secreted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit polypeptide. More particularly, the targeted $\alpha_2\delta$ -2, $\alpha_2\delta$ -3 or $\alpha_2\delta$ -4 subunit binding site is preferably the [^3H]gabapentin binding site. The various
35 parameters of the method of the invention are described in further detail below.

EXAMPLES

5 Example 1

Construction of a nucleotide sequence encoding a soluble secreted human $\alpha_2\delta$ -2 subunit polypeptide deletion mutant of SEQ ID N°23

a) Primer design

- 10 PCR primers were designed to generate the secreted soluble human $\alpha_2\delta$ -2 deletion mutant of SEQ ID N° 23 as follows:

5' PCR primer: This was designed to engineer in a KOZAK translation initiation consensus sequence prior to the coding sequence (Kozak *JBC* 266 19867-19870)

- 3' PCR primer: This was designed to engineer in six histidine residues followed by a stop-codon at the desired location in the coding sequence. In addition to the stop codon the $\alpha_2\delta$ -2 primers also included an *Eco* RI restriction site.

The bold region in each primer sequence denotes the 'tagged' region; addition of sequences not present in the template. Primers were custom synthesized by Perkin Elmer Applied Biosystems UK to the ABI ready pure grade, supplied lyophilized then resuspended to 15 μ M in 10mM TE. JB197 and 198 were provided with 5' phosphate groups:

5' Primer JB197 (5'-TCGCCACCATGGCGGTGCCGGCTC-3', SEQ ID N°25)

3' Primer JB198 (5'-TCGGAATTCCTCAGTGATGGTGATGGTGATGGGCCCGCGGCCACAGTC-3', SEQ ID N°28)

b) Protocol for PCR mediated 5' Kozak and 3' 6His tagging of human $\alpha_2\delta$ -2

30

The full length human $\alpha_2\delta$ -2 gene (Gen Bank Accession Number AF042792) in a pcDNA 3 vector as described in Brown, J.P. and Gee, N.S., Cloning and deletion mutagenesis of the $\alpha_2\delta$ calcium channel subunit from porcine cerebral cortex, *The journal of biological chemistry*, 273(39):25458-25465) was used as the template in the following PCR reaction.

- 35 The reagents were added in the following order in triplicate to a 96 well PCR plate:

	μ l
10x Pfx Amplification buffer	5
10mM dNTPs	1.5

	50mM MgSO ₄	1
	15μM JB197	1.5
	15μM JB198	1.5
	100ng/μl pcDNA3.1-humans-α ₂ δ-2	1
5	10x PCR Enhancer	5
	H ₂ O	32.7
	2.5 UNITS/μL PFX POLYMERASE	0.8μL

The plate was the cycled on an MJ Tetrad DNA engine according to the following cycling
10 conditions:

94°C / 2mins

followed by:

for 30 cycles 94°C / 45sec

15 58°C / 45sec

68°C / 4mins

followed by:

68°C / 10mins

followed by:

20 hold at 4°C

The 3366bp product was then gel purified from a 1% TAE agarose gel using QIAEX beads and eluted in approximately 50μl TE.

25

Example 2

Cloning of the PCR fragments of Example 1 into the Baculovirus transfer vector pFastBac1

30 The PCR products of Example 1 were cloned into *Sfu* I digested, calf intestinal phosphatase dephosphorylated, phenol chloroform extracted and QIAEX gel purified pFastBac1 (Life Technologies) using the Rapid DNA ligation kit (Roche Diagnostics) transforming XL1-blue (α₂δ-1b) *E. Coli* cells:

a) Screening for positive recombinants

- Given that the PCR product was cloned by blunt-end ligation a screen was required to select a recombinant with the gene ligated in the positive orientation with respect to the polyhedrin promoter in pFastBac1. This was achieved by restriction digest of miniprep DNA (Qiagen miniprep kit) prepared from colony minicultures and analysis on a 1% TAE agarose gel. A positive clone was identified according to the following digest patterns:

SEQ ID N° 23 in pFastBac1

Eco RI digest performed on miniprep DNA

Predicted fragments (bp)

PCR product cloned in a positive orientation

4773 and 3368

PCR product cloned in a negative orientation

8127 and 14

15 b) Sequencing analysis of selected clones

One positive was selected for this clone and used to prepare a plasmid DNA stock of the desired construct (QIAGEN maxi kit). Confirmatory sequence reactions were performed using the Big Dye terminator sequencing kit and run on an ABI 310 Prism Genetic Analyzer. Sequence analysis of both coding strands was performed using a selection of sequencing oligonucleotide primers.

Example 3

Protocol for establishing baculovirus banks for the expression of the $\alpha_2\delta$ -2 deletion mutant SEQ ID N°23

Essentially, the protocol used to generate the baculovirus banks is that outlined in the Life Technologies Bac-to Bac™ baculovirus expression systems manual.

a) Transposition of DH10Bac *E Coli* cells

- One ng (5µl) of the recombinant pFastBac-1 construct containing the nucleotide sequence encoding the porcine $\alpha_2\delta$ -2 deletion mutant of SEQ ID N°23 was added to 100µl of DH10Bac cells thawed on ice. The cells were then mixed gently by tapping the tube then incubated on ice for 30minutes before heat shock treatment by incubation in a 42°C water bath for 45 seconds. The mixture was then chilled on ice for 2 minutes before the addition of 900µl of S.O.C. medium. The mixture was then placed in a shaking incubator (200rpm) at 37°C for 4hours. The cells were then serially diluted (10 fold dilutions from 10⁻¹ to 10⁻³)

and 10µl of each dilution plated on LB agar plates containing 50µg/ml kanamycin, 7µg/ml gentamicin, 10µg/ml tetracycline, 100µg/ml Bluo-gal and 40µg/ml IPTG. The plates were incubated at 37°C for between 1 and 3 days until discrete colonies of blue and white colour were discernible.

5

b) Isolation of recombinant DNA

White colonies (containing the recombinant bacmid) were picked and grown for 24 hours (to stationary phase) at 37°C with shaking (200rpm) in 2ml of LB containing 50µg/ml kanamycin, 7µg/ml gentamicin and 10µg/ml tetracycline. 1.5ml of culture was then

10 transferred to a microfuge tube and centrifuged at 14,000xg for 1 minute. The supernatant was removed and the cells resuspended gently in 0.3ml of 15mM Tris-HCl (pH8.0), 10mM EDTA, 100µg/ml RNase A. 0.3ml of 0.2N NaOH, 1% SDS was then added and the mixture mixed gently before incubation at 22°C for 5 minutes. Then 0.3ml of 3M

15 Potassium acetate (pH5.5) was added and the sample placed on ice for 10 minutes. After centrifugation at 14,000xg for 10 minutes the supernatant was transferred to a tube containing 0.8ml of isopropanol, mixed then placed on ice for 10 minutes before centrifugation at 14,000xg for 10 minutes. The supernatant was then discarded and the pellet rinsed with 0.5ml of 70% ethanol before centrifugation at 14,000xg for 5 minutes. This 70% ethanol rinse was then repeated before removing all of the supernatant and air

20 drying the pellet for 10 minutes at room temperature. The pellet was finally resuspended in 40µl of TE.

c) Transfection of sf9 cells with the recombinant bacmid DNA

A 6-well tissue culture plate was seeded with 0.9×10^6 sf9 cells (cells at log phase having

25 grown from a culture passaged at 0.3×10^6 cells/ml) per 35mm well in 2ml of Sf-900 II SFM media containing 50units/ml penicillin and 50µg/ml streptomycin. Cells were left to attach at 27°C for 1 hour. Bacmid DNA prepared as described above (5µl) was added to 200µl of Sf-900 II SFM media containing 6µl of CELLFECTIN and mixed before incubation at room temperature for 45 minutes. The cells were washed once with 2ml of

30 Sf-900 II SFM media without antibiotics then 0.8ml of Sf-900 II SFM media was added to each tube containing the lipid-DNA complex. The wash buffer was removed from the cells and the 1ml of diluted lipid-DNA complex overlaid on the cells. The cells were incubated for 5 hours at 27°C after which time the transfection mixture was removed and 2ml of Sf-

900 II SFM media containing 50units/ml penicillin and 50µg/ml streptomycin was added. The cells were then incubated for 72 hours.

After incubation for 72 hours the media was removed from the cells and centrifuged at 500xg for 5 minutes. The supernatant was then transferred to a fresh tube, this was labelled as the P0 bank and stored at 4°C in the dark. The P1 bank was prepared by passaging sf9 cells at approx 5×10^6 cells/ml to 2×10^6 cells/ml (100ml in a 250ml Erlenmeyer flask) and adding 0.5ml of the P0 bank harvested above. The cells were then incubated shaking (200rpm) at 27°C for 4 days. Under sterile conditions the culture was centrifuged at 500xg for 10 minutes and the supernatant 0.2µM filtered (P1 bank). The P2 bank was prepared by adding 2ml of P1 bank per 400ml culture (in 1L Erlenmeyer flasks) passaged as above to 2×10^6 cells/ml. The culture was incubated as before for 4 days and the supernatant harvested and filtered as described for the P1 bank. The supernatant was first pooled then aliquoted (10ml) and stored at 4°C.

15

Example 4
Expression of the $\alpha_2\delta$ -2 deletion mutant of SEQ ID N°23

To sf9 cells passaged from $\sim 5 \times 10^6$ cells/ml to 2×10^6 cells/ml in Sf-900 II SFM media was added 0.1ml virus per 100ml of cells of the appropriate viral bank (400ml volumes in 1L Erlenmeyer flasks). The cells were then cultured for 4-5 days at 27°C with 110rpm shaking. Expression of the protein was confirmed by SDS-PAGE and Western blotting using an anti penta-His monoclonal antibody (Qiagen) and was detected in the culture supernatant and cell lysate.

25

Example 5
Purification of $\alpha_2\delta$ -2deletion mutant of SEQ ID N°23

The- $\alpha_2\delta$ -2 deletion mutant of SEQ ID N°23 was purified from the cell lysate following the purification strategy outlined below:

30

The culture was centrifuged at 6,000xg for 10 minutes and the supernatant removed. The weight of the cell pellet was determined before re-suspension in 20mM Tris pH8.0, 100mMKCl, 1% P40-Nonidet (100ml per 20g of wet cells). A protease inhibitor cocktail

(Sigma, Cat# P8849), 1ml/L, was added to the mixture. The solution was then stirred for 10 minutes before centrifugation for 1 hour at 30,000xg and 4°C. The supernatant was concentrated (30kDa cut off) to approx. ~300ml then centrifuged for 1 hour at 100,000xg.

- 5 Supernatant containing the soluble proteins was diluted 1:3 in 10mM Tris-HCl pH8.0 (equilibration buffer) and loaded onto a pre-equilibrated Q-Sepharose column (2.5cm i.d. x 30cm h.) at a flow rate of 900ml/h. After washing with equilibration buffer until a stable A_{280nm} baseline had been achieved, protein was eluted with 20mM Tris-HCl pH8.0, 0.5M KCl, 10mM Imidazole.

10

The eluate was then loaded onto a Ni-NTA (Qiagen) column (2.5cm i.d. x 6cm h.) pre-equilibrated in 20mM Tris pH8.0, 0.5M KCl, 10mM Imidazole at a flow rate of 2 ml/min.

The column was washed successively with buffer A (20mM Tris pH8.0, 0.5M KCl, 20mM Imidazole), buffer B (100mM Tris-HCl pH8.0, 1M KCl), and buffer A again. Elution was

15

performed with buffer C (20mM Tris-HCl pH8.0, 100mM KCl, 0.5M Imidazole). The Ni-NTA eluate (~50ml) was concentrated (30kDa cut-off) to ~2ml and applied at 1ml/min and in 0.2ml aliquots, to an FPLC Superdex-200 column equilibrated in 10mM HEPES, pH7.4, 150mM NaCl. Fractions containing the polypeptide of SEQ ID N°23 were pulled.

20

Example 6

SPA assay of [3 H]gabapentin binding to the secreted soluble human $\alpha_2\delta$ -2 subunit of SEQ ID N°23

- 25 The assay is carried out at 21°C. Assay components are added in the following order (all reagents are diluted in 10mM HEPES (pH 7.4 at 21°C) to 96-well Optiplates:

25µl imidazole at various concentrations (diluted from a 1M stock pH8.0, see assay details)

50µl 10mM HEPES pH 7.4

30

25µl (50mg) SPA beads (Amersham)

100µl s- $\alpha_2\delta$ -2 subunit polypeptide of SEQ ID No 23 (2µl protein

diluted to 100µl)

25µl radioligand ([3 H]gabapentin obtained from example 5

Immediately after adding radioligand, the optiplates were loaded in the Packard Top Count scintillation counter to follow the binding time course. Imidazole was first used in the assay to optimize the specific interaction of the protein's 6His tag with the SPA bead. Imidazole itself (up to 100mM) in the filtration assay has no effect on [³H]gabapentin

5 binding (n=1).

Example 7

Ni Flashplate assay of [³H]gabapentin binding to secreted soluble human $\alpha_2\delta$ -2 (SEQ ID N°23)

10 Assays are carried out at 21°C in a final volume of 250µl in 96-well NEN Ni chelate flash plates. Assay components are added in the following order (all reagents were diluted in 10mM HEPES (pH 7.4 at 21°C)):

25µl 10mM HEPES pH7.4

25µl imidazole at various concentrations (diluted from a 1M stock
15 pH8.0, see assay details)

75µl 10mM HEPES pH 7.4

100µl s- $\alpha_2\delta$ -2-6His (2µl protein diluted to 100µl) obtained from
example 5

25µl radioligand ([³H]gabapentin (65Ci/mmole)

20 Immediately after adding the radioligand, flash plates are loaded in the Packard Top Count scintillation counter to follow the binding time course. The '[³H] flash plate' programme (cpm) is used to monitor activity. Imidazole is first used in the assay to optimize the specific interaction of the protein's 6His tag with the Ni flashplate.

Example 8

Ni Flashplate assay of [³H]Leucine binding to secreted soluble human $\alpha_2\delta$ -2-6His

30 The procedure described in example 7 is repeated, except that [³H]gabapentin is replaced by 25 ul (10.1 nM) of [³H]Leucine (141 Ci/mmole).

1233

sub
B¹²

- ~~S pH 7.4~~
~~concentra~~

GraphPad Prism software is used to process competition curve data and determine IC_{50} and hill slope values. Twelve point competition curves with half log dilution steps of test compounds are used in the experiments.

~~Filter binding assay of [³H]gabapentin binding to the recombinant polypeptide of~~
~~SEQ ID N°23~~

10

200µl Polypeptide of SEQ ID N°23 (3µl protein diluted to 200µl)

25 μ l radioligand ($[^3\text{H}]$ gabapentin (65Ci/mmol))

15

Sub
B³ 25

20

25

25

30

30

The bold region in each primer sequence denotes the ‘tagged’ region; addition of sequences not present in the template. Primers were custom synthesized by Perkin Elmer Applied Biosystems UK to the ABI ready pure grade, supplied lyophilized then

resuspended to 15 μ M in 10mM TE. JB201 and 202 were provided with 5' phosphate groups:

5' Primer JB201 (5'-TCGCCACCATGGCCGGGCCGGGC-3', SEQ ID N°27)

3' Primer JB202 (5'-TCTCAGTGATGGTGTATGGTGTATGCGATGCACCCCCACACTCTC-3', SEQ ID N°28)

b) Protocol for PCR mediated 5' Kozak and 3' 6His tagging of mouse $\alpha_2\delta$ -3

- 10 The full length mouse $\alpha_2\delta$ -3 gene (Gen Bank Accession number AJ010949) in the pcDNA3 vector as described in Brown, J.P. and Gee, N.S., Cloning and deletion mutagenesis of the $\alpha_2\delta$ calcium channel subunit from porcine cerebral cortex, *The journal of biological chemistry*, 273(39):25458-25465) was used as the template in the following PCR reaction.

15

The reagents were added in the following order in triplicate to a 96 well PCR plate:

	μ l
10x Pfx Amplification buffer	5
10mM dNTPs	1.5
20 50mM MgSO ₄	1
15 μ M JB201	1.5
15 μ M JB202	1.5
100ng/ μ l pcDNA3-mouse- $\alpha_2\delta$ -3	1
10x PCR Enhancer	5
25 H ₂ O	32.7
2.5 UNITS/ μ L PFX POLYMERASE	0.8 μ L

The plate was the cycled on an MJ Tetrad DNA engine according to the following cycling conditions:

30

94°C / 2mins

followed by:

for 30 cycles 94°C / 45sec

60°C / 45sec

68°C / 4mins

followed by:

68°C / 10mins

followed by:

5 hold at 4°C

The 3244bp product was then gel purified from a 1% TAE agarose gel using QIAEX beads and eluted in approximately 50µl.

The truncated protein of SEQ ID N°24 was expressed such the procedure of example 2,3
10 and 4.

REFERENCES

- Perez-Reyes, E., and Schneider, T. (1994) *Drug Dev. Res.* **33**, 295-318
- Catterall, W. A. (1995) *Annu. Rev. Biochem.* **64**, 493-531
- 15 - Birnbaumer, L., Campbell, K. P., Catterall, W. A., Harpold, M. M., Hofmann, F., Horne, W. A., Mori, Y., Schwartz, A., Snutch, T. P., Tanabe, T., and Tsien, R. W. (1994) *Neuron* **13**, 505-506
- Brust, P. F., Simerson, S., McCue, A. F., Deal, C. R., Schoonmaker, S., Williams, M. E., Velicelebi, G., Johnson, E. C., Harpold, M. M., and Ellis, S. B. (1993)
- 20 *Neuropharmacology* **32**, 1089-1102
- Itagaki, K., Koch, W. J., Bodi, L., Klockner, U., Shish, D. F., and Schwartz, A. (1992) *FEBS Lett.* **297**, 221-225
- Mikami, A., Imoto, F., Tanabe, T., Niidome, T., Mori, Y., Takeshima, H., Narumiya, S., and Numa, S. (1989) *Nature* **340**, 230-233
- 25 - Mori, Y., Friedrich, T., Kim, M. S., Mikami, A., Nakai, J., Ruth, P., Bosse, E., Hofmann, F., Flockerzi, V., Furuichi, T., Mikoshiba, K., Imoto, K., Tanabe, T., and Numa, S. (1991) *Nature* **350**, 398-402
- Singer, D., Biel, M., Lotan, I., Flockerzi, V., Hofmann, F., and Dascal, N. (1991) *Science* **253**, 1553-1657
- 30 - Ramsay, R. E. (1994) *Neurology* **44**, Suppl. 5, 23-30
- Watson, W. P., and Little, H. J. (1995) *Br. J. Pharmacol.* **116**, 33P (abstr.)
- Singh, L., Field, M. J., Ferris, P., Hunter, J. C., Oles, R. J., Williams, R. G., and Woodruff, G. N. (1996) *Psychopharmacology* **127**, 1-9
- Xiao, W. H., and Bennet, G. L. (1995) *Soc. Neurosci.* **21**, 897 (abstr.)
- 35 - Mellick, G. A., Mellicy, L. B., and Mellick, L. B. (1995) *J. Pain Symptom Manage.* **10**, 265-266

09297550-091690

- Shimoyama, N., Shimoyama, M., Davis, A. M., Inturrisi, C. E., and Elliott, K. J. (1997) *Neurosci. Lett.* **222**, 65-67
- Segal A. Z., and Rordorf, G. (1996) *Neurology* **46**, 1175-1176
- Mellick, G. A., and Mellick, L. B. (1996) *Sleep* **19**, 224 -226
- 5 - Patel, J., and Naritoku, D. K (1996) *Clin. Neuropharmacol.* **19**,185-188
- Suman Chauhan, N., Webdale, L., Hill, D. R., and Woodruff, G. N. (1993) *Eur. J. Pharmacol.* **244**, 293-301
- Macdonald, R. L., and Kelly, F. M. (1993) *Epilepsia* **34**, Suppl. 5, S1-S8
- Taylor, C. P. (1994) *Neurology* **44**, Suppl. 5, 10 -16
- 10 - Gotz, E., Feuerstein, T. J., Lais, A., and Meyer, D. K (1993) *Arzneimittelforschung* **43**, 636-638
- Loscher, W., Honack, D., and Taylor, C. P. (1991) *Neurosci. Lett.* **128**,150-154
- Honmou, O., Knesis, J. D., and Richerson, G. B. (1995) *Epilepsy Res.* **20**, 193-202
- Honmou, O., Oyelese, A. A., and Kocsis, J. D. (1995) *Brain Res.* **692**,273-277
- 15 - Petroff, O. A. C., Rothman, D. L., Behar, K. L., Lamoureux, D., and Mattson, R. H. (1996) *Ann. Neurol.* **39**, 95-99
- Reimann, W. (1983) *Eur. J. Pharmacol.* **94**, 341-344
- Dooley, D. J., Bartoszyk, G. D., Hartenstein, J., Reimann, W., Rock, D. M., and Satzinger, G. (1986) *Golden Jubilee Conference and Northern European Epilepsy Meeting*. Abstracts, University of York, UK, September 1986 (Abstract 8).
- 20 - Thurlow, R. J., Brown, J. P., Gee, N. S., Hill, D. R., and Woodruff, G. N. (1993) *Eur. J. Pharmacol.* **247**,341-345
- Gee, N. S., Brown, J. P., Dissanayake, V. U. I., Offord, J., Thurlow, R., and Woodruff, G. N. (1996) *J. Biol. Chem.* **271**, 5768-5776
- 25 - Dissanayake, V. U. I., Gee, N. S., Brown, J. P., and Woodruff, G. N. (1997) *Br. J. Pharmacol.* **120**, 833-840
- Taylor, C. P., Vartanian, M. G., Yuen, P. W., Bigge, C., Suman Chauhan, N., and Hill, D. R. (1993) *Epilepsy Res.* **14**,11-15
- Rock, D. M., Kelly, K. M., and Macdonald, R. L. (1993) *Epilepsy Res.* **16**, 89 -98
- 30 - Wamil, A. W., Mclean, M. J., Nashville, T. N., and Taylor, C. P. (1991) *Neurology* **41**, Suppl. 1, 140 (abstr.)
- De Jongh, K S., Warner, C., and Catterall, W. A. (1990) *J. Biol. Chem.* **265**, 14738-14741
- Jay, S. D., Sharp, A. H., Kahl, S. D., Vedvick, T. S., Harpold, M. M., and Campbell, K. P. (1991) *J. Biol. Chem.* **266**, 3287-3293
- 35 - Burgess, A. J., and Norman, R. I. (1988) *Eur. J. Biochem.* **178**, 527-533
- Ellis, S. B., Williams, M. E., Ways, N. R., Brenner, R., Sharp, A. H., Leung, A. T., Campbell, K. P., McKenna, E., Koch, W. J., Hai, A., Schwartz, A., and Harpold, M. M. (1988) *Science* **241**, 1661-1664

- Brickley, K., Campbell, V., Berrow, N., Leach, R., Norman, R. I., Wray, D., Dolphin, A. C., and Baldwin, S. A- (1995) *FEBS Lett.* **364**,129-133
- Brice, N. L., Berrow, N. S., Campbell, V., Page, K. M., Brickley, K., Tedder, I., Dolphin, A C. (1997) *Eur. J. Neurosci.* **9**, 749-759
- 5 - Wiser, O., Trus, M., Tobi, D., Halevi, S., Giladi, E., and Atlas, D. (1996) *FEBS Lett.* **379**,15-20
- Xu, X., and Arnason, U. (1994) *Gene (Amst.)* **148**, 357-362
- Williams, M. E., Feldman, D. H., McCue, A. F., Brenner, R., Velicelebi, G., Ellis, S. B., and Harpold, M. M. (1992) *Neuron* **8**, 71-84
- 10 - Kim, H. L., Kim, H., Lee, P., King, R. G., and Chin, H. (1992) *Proc. Natl. Acad Sci. U.S.A.* **89**,3251-3255
- Brown, J. P., Dissanayake, V. U. K., Briggs, A. R., Milic, M. R., and Gee, N. S. (1998) *Anal. Biochem.* **255**, 236-243
- Higuchi, R. (1990) in *PCR Protocols: A Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J. eds) pp. 177-183, Academic Press, Ltd., London
- 15 - Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248-252
- Kyte, J., and Doolittle, F. (1982) *J. Mol. Biol.* **157**, 105-132
- Summers, M. F., Henderson, L. E., Chance, M. R., Bess, J. W., Jr., South, T. L., Blake, P. R., Sagi, I., Perez-Alvarado, G., Sowder, R. C., Hare, D. R., and Arthur, L. O. (1992) *Protein Sci.* **1**, 563-574
- 20 - Klug, A. and Rhodes, D. (1987) *Trends. Biochem. Sci.* **12**, 464-469
- Pieler, T., and Bellefroid, E. (1994) *Mol. Biol. Rep.* **20**, 1-8
- Preston, R. A., Manolson, M. F., Becherer, M., Weidenhammer, E., Kirkpatrick, D., Wright, R., and Jones, E. W. (1991) *Mol. Cell, Biol.* **11**, 5801-5812
- 25 - Tan, X., Waterham, H. R., Veenhuis, M., and Cregg, J. M. (1995) *J. Cell Biol.* **128**,307-319
- Scotland, P. B., Colledge, M., Melnikova, I., Dai, Z., and Froehner, S. C. (1993) *J. Cell Biol.* **123**, 719-728
- 30 - Henderson, L. E., Copeland, T. D., Sowder, R. C., Smythers, G. W., and Oroszlan, S. (1981) *J. Biol. Chem.* **256**, 8400- 8406
- Beaucage et al., *Tetrahedron Lett* (1981) **22**: 1859-1862.
- Brown El., Belagaje R, Ryan MJ, Khorana HG, *Methods Enzymol* (1979); **68**, 109-151.
- Feldman and Steg, (1996) *Medecine/Sciences, synthese*, **12**, 47-55.
- 35 - Houbenweyl, (1974), in *Meuthode der Organischen Chemie*, E. Wunsch Ed., Volume 15-I et 15-II, Thieme, Stuttgart.
- Koch Y. (1977), *Biochem. Biophys. Res. Commun.*, **74**, 488-491.
- Kohler G. and Milstein C., (1975) *Nature*, **256**, 495.

- Kozbor et al., (1983) *Hybridoma*, **2**(1), 7-16.
- Leger OJ, et al. (1997) *Hum Antibodies* , **8**(1), 3-16.
- Martineau P, Jones P, Winter G. (1998), *J. Mol Biol*, **280**(1), 117-127.
- Merrifield RB, 1965a, *Nature*, **207**(996), 522-523.
- 5 - Merrifield RB, 1965b, *Nature*, **207** (996), 22-523.
- Narang SA, Hsiung HM, Brousseau R, *Methods Enzymol* 1979, **68**, 90-98.
- Ohno et al., (1994), *Science*, **265**, 781-784.
- O'Reilly et al., (1992) *Baculovirus expression vectors: a Laboratory Manual*. W.H. Freeman and Co., New York.
- 10 - Ridder R. Schmitz R, Legay F, Gram H, (1995) *Biotechnology* (NY), **13**(3), 255-260.
- Smith et al., (1983), *Mol. Cell. Biol.*, **3**, 2156-2165.
- Sternberg N.L. (1992), *Trends Genet*, **8**, 1-16.
- Sternberg N.L. (1994) *Mamm. Genome*, **5**, 397-404.
- Sambrook, J. Fritsch, E.F. and T. Maniatis (1989). *Molecular cloning: a laboratory*
- 15 *manual*, 2ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Sanchez-Pescador R., (1988), *J. Clin. Microbiol.*, **26**(10), 1934-1938.
- Urdea et al., MS (1988) *Nucleic Acids Research*, **11**, 4937-4957.
- Urdea et al., MS (1991) *Nucleic Acids Symp Ser.*, **24**, 197-200.

Accepted for publication

SEQUENCE LISTING**1 - human nucleotide sequence**

ATGGCGGTGCCGGCTCGGACCTGCGGCGCCTCTCGGCCCCGGCCCAGCGCGGACTGCGCGCC
 5 CCTGGCCCCGGCTGCGGCCCCCACCCTGGCCCCGGCACCCGGCGCCCGACGTCCGGGGCCCCC
 GCGCCCCGCTGTGGCTGCTGC&TGCCGCTTCTACCGCTGCTCGCCGCCCCCGGCGCCTCTGC
 CTACAGCTTCCCCCAGCAGCACACGATGCAGCACTGGGCCCCGGCGTCTGGAGCAGGAGGTC
 GACGGCGTGATGCGGATTTTGGAGGCGTCCAGCAGCTCCGTGAGATTTACAAGGACAACC
 GGAACCTGTTTCGAGGTACAGGAGAATGAGCCTCAGAAGTTGGTGGAGAAGGTGGCAGGGGA
 10 CATTGAGAGCCTTCTGGACAGGAAGGTGCAGGCCCTGAAGAGACTGGCTGATGCTGCAGAG
 AACTTCCAGAAAGCACACCGCTGGCAGGACAACATCAAGGAGGAAGACATCGTGACTATG
 ACGCCAAGGCTGACGCTGAGCTGGACGACCCTGAGAGTGAGGATGTGGAAAGGGGGTCTAA
 GGCCAGCACCCCTAAGGCTGGACTTCATCGAGGACCCAACTTCAAGAACAAGGTCAACTAT
 TCATACGCGGCTGTACAGATCCCTACGGACATCTACAAAGGCTCCACTGTCATCCTCAATG
 15 AGCTCAACTGGACAGAGGCCCTGGAGAATGTGTTTCATGGAAAACCGCAGACAAGACCCAC
 ACTGCTGTGGCAGGTCTTCGGCAGCGCCACAGGAGTCACTCGCTACTACCCGGCCACCCCG
 TGGCGAGCCCCCAAGAAGATCGACCTGTACGATGTCCGAAGGAGACCCTGGTATATCCAGG
 GGGCCTCGTCACCCAAAGACATGGTCATCATCGTGGATGTGAGTGGCAGTGTGAGCGGCCT
 GACCCTGAAGCTGATGAAGACATCTGTCTGCGAGATGCTGGACACGCTGTCTGATGATGAC
 20 TATGTGAATGTGGCCTCGTTCAACGAGAAGGCACAGCCTGTGTCATGCTTCACACACCTGG
 TGCAGGCCAATGTGCGCAACAAGAAGGTGTTCAAGGAAGCTGTGCAGGGCATGGTGGCCAA
 GGGCACACAGGCTACAAGGCCGGCTTTGAGTATGCCTTTGACCAGCTGCAGAACTCCAAC
 ATCACTCGGGCCAACTGCAACAAGATGATCATGATGTTACGGATGGTGGTGAGGACCGCG
 TGCAGGACGTCTTTGAGAAGTACAATTGGCCAAACCGGACGGTGCGCGTGTTTACTTTCTC
 25 CGTGGGGCAGCATAACTATGACGTACACCGCTGCAGTGGATGGCCTGTGCCAACAAAGGC
 TACTATTTTGAGATCCCTTCCATCGGAGCCATCCGCATCAACACACAGGAATATCTAGATG
 TGTTGGGCAGGCCCATGGTGCTGGCAGGCAAGGAGGCCAAGCAGGTTCAGTGGACCAACGT
 GTATGAGGATGCACTGGGACTGGGGTTGGTGGTAACAGGGACCCTCCCTGTTTTCAACCTG
 ACACAGGATGGCCCTGGGGAAAAGAAGAACAGCTGATCCTGGGCGTGATGGGCATTGACG
 30 TGGCTCTGAATGACATCAAGAGGCTGACCCCCAACTACACGCTTGGAGCCAACGGCTATGT
 GTTTGCCATTGACCTGAACGGCTACGTGTTGCTGCACCCCAATCTCAAGCCCCAGACCACC
 AACTTCCGGGAGCCTGTGACTCTGGACTTCCTGGATGCGGAGCTAGAGGATGAGAACAAGG
 AAGAGATCCGTGCGAGCATGATTGATGGCAACAAGGGCCACAAGCAGATCAGAACGTTGGT
 CAAGTCCCTGGATGAGAGGTACATAGATGAGGTGACACGGAACCTACACCTGGGTGCCTATA
 35 AGGAGCACTAACTACAGCCTGGGGCTGGTGCTCCACCCCTACAGCACCTTCTACCTCCAAG
 CCAATCTCAGTGACCAGATCCTGCAGGTCAAGTATTTTGAAGTTCCTGCTCCCCAGCAGCTT
 TGAGTCTGAAGGACACGTTTTTCATTGCTCCCAGAGAGTACTGCAAGGACCTGAATGCCTCA
 GACAACAACACCGAGTTCCTGAAAACTTTATTGAGCTCATGGAGAAAGTGACTCCAGACT
 CCAAGCAGTGCAACAACCTTCCTTCTGCACAACCTGATCTTGGACACGGGCATCACGCAGCA

GCTGGTAGAGCGTGTGTGGAGGGACCAGGATCTCAACACGTACAGCCTACTGGCCGTGTTT
 GCTGCCACAGACGGTGGCATCACCCGAGTCTTCCCCAACAAGGCAGCTGAGGACTGGACAG
 AGAACCCCTGAGCCCTTCAATGCCAGCTTCTACCGCCGACGCTGGATAACCACGGTTATGT
 CTTCAAGCCCCCACACCAGGATGCCCTGTAAAGGCCGCTGGAGCTGGAGAATGACACTGTG
 5 GGCATCCTCGTCAGCACAGCTGTGGAGCTCAGCCTAGGCAGGCGCACACTGAGGCCAGCAG
 TGGTGGGCGTCAAGCTGGACCTAGAGGCTTGGGCTGAGAAGTTCAAGGTGCTAGCCAGCAA
 CCGTACCCACCAAGACCAGCCTCAGAAGTGC GGCCCCAACAGCCACTGTGAGATGGACTGC
 GAGGTTAACAATGAGGACTTACTCTGTGTCTCATTTGATGATGGAGGATTCTCTGGTGTCTGT
 CAAACCAGAACCATCAGTGGGACCAGGTGGGCAGGTTCTTCAGTGAGGTGGATGCCAACCT
 10 GATGCTGGCACTCTACAATAACTCCTTCTACACCCGCAAGGAGTCTATGACTATCAGGCA
 GCCTGTGCCCCCTCAGCCCCCTGGCAACCTGGGTGCTGCACCCCGGGGTGTCTTTGTGCCCA
 CCGTTGCAGATTTCTTAACCTGGCCTGGTGGACCTCTGCTGCCGCTGGTCCCTGTTCCA
 GCAGCTTCTCTACGGCCTCATCTACCACAGCTGGTTCCAAGCAGACCCCGCGGAGGCCGAG
 GGGAGCCCCGAGACGCGCGAGAGCAGCTGCGTCATGAAACAGACCCAGTACTACTTCGGCT
 15 CGGTAAACGCCTCTACAACGCCATCATCGACTGCGGAAACTGCTCCAGGCTGTTCCACGC
 GCAGAGACTGACCAACACCAATCTTCTCTTTGTGGTGGCCGAGAAGCCGCTGTGCAGCCAG
 TGCGAGGCTGGCCGG

2 - human nucleotide sequence

20 ATGGCGGTGCCGGCTCGGACCTGCGGCGCCTCTCGGCCCCGGCCAGCGCGGACTGCGCGCC
 CCTGGCCCCGCTGCGGCCCCCACCCTGGCCCCGGCACCCGGCGCCCGACGTCCGGGCCCCC
 GCGCCCGCTGTGGCTGCTGCTGCCGCTTCTACCGCTGCTCGCCGCCCCCGCGCCTCTGCC
 TACAGCTTCCCCCAGCAGCACACGATGCAGCACTGGGCCCCGGCGTCTGGAGCAGGAGGTCTG
 ACGGCGTGATGCGGATTTTTTGGAGGCGTCCAGCAGCTCCGTGAGATTTACAAGGACAACCG
 25 GAACCTGTTTCAGGTACAGGAGAATGAGCCTCAGAAGTTGGTGGAGAAGGTGGCAGGGGAC
 ATTGAGAGCCTTCTGGACAGGAAGGTGCAGGCCCTGAAGAGACTGGCTGATGCTGCAGAGA
 ACTTCCAGAAAGCACACCGCTGGCAGGACAACATCAAGGAGGAAGACATCGTGTACTATGA
 CGCCAAGGCTGACGCTGAGCTGGACGACCCTGAGAGTGAGGATGTGGAAAGGGGGTCTAAG
 GCCAGCACCTAAGGCTGGACTTCATCGAGGACCCAACTTCAAGAACAAAGGTCAACTATT
 30 CATACGCGGCTGTACAGATCCCTACGGACATCTACAAAGGCTCCACTGTCATCCTCAATGA
 GCTCAACTGGACAGAGGCCCTGGAGAATGTGTTTATGGAAAACCGCAGACAAGACCCACA
 CTGCTGTGGCAGGTCTTCGGCAGCGCCACAGGAGTCACTCGCTACTACCCGGCCACCCGCT
 GGCGAGCCCCCAAGAAGATCGACCTGTACGATGTCCGAAGGAGACCTGGTATATCCAGGG
 GGCCTCGTCACCCAAAGACATGGTCATCATCGTGGATGTGAGTGGCAGTGTGAGCGGCCTG
 35 ACCCTGAAGCTGATGAAGACATCTGTCTGCGAGATGCTGGACACGCTGTCTGATGATGACT
 ATGTGAATGTGGCCTCGTTCAACGAGAAGGCACAGCCTGTGTATGCTTACACACCTGGT
 GCAGGCCAATGTGCGCAACAAGAAGGTGTTCAAGGAAGCTGTGCAGGGCATGGTGGCCAAG
 GGCACCACAGGCTACAAGGCCGGCTTTGAGTATGCCTTTGACCAGCTGCAGAACTCCAACA

09267550-091600

TCACTCGGGCCAACTGCAACAAGATGATCATGATGTTCCAGGATGGTGGTGAGGACCGCGT
GCAGGACGTCTTTGAGAAGTACAATTGGCCAAACCGGACGGTGC GCGTGT TTTACTTTCTCC
GTGGGGCAGCATAACTATGACGTCAACCCGCTGCAGTG GATGGCCTGTGCCAACAAAGGCT
ACTATTTTGAGATCCCTTCCATCGGAGCCATCCGCATCAACACACAGGAATATCTAGATGT
5 GTTGGGCAGGCCCATGGTGCTGGCAGGCAAGGAGGCCAAGCAGGTT CAGTGGACCAACGTG
TATGAGGATGCACTGGGACTGGGGTTGGTG GTAACAGGGACCCTCCCTGTTTTC AACCTGA
CACAGGATGGCCCTGGGGAAAAGAAGAACCAGCTGATCCTGGGCGTGATGGGCATTGACGT
GGCTCTGAATGACATCAAGAGGCTGACCCCCAACTACACGCTTGGAGCCAACGGCTATGTG
TTTGCCATTGACCTGAACGGCTACGTGTTGCTGCACCCCAATCTCAAGCCCCAGACCACCA
10 ACTTCCGGGAGCCTGTGACTCTGGACTTCCTGGATGCGGAGCTAGAGGATGAGAACAAAGGA
AGAGATCCGTCGGAGCATGATTGATGGCAACAAGGGCCACAAGCAGATCAGAACGTTGGTC
AAGTCCCTGGATGAGAGGTACATAGATGAGGTGACACGGAACTACACCTGGGTGCCTATAA
GGAGCACTAACTACAGCCTGGGGCTGGTGCTCCCAACCCTACAGCACCTTCTACCTCCAAGC
CAATCTCAGTGACCAGATCCTGCAGGTCAAGTATTTTGAGTTCCTGCTCCCCAGCAGCTTT
15 GAGTCTGAAGGACACGTTTTTCATTGCTCC CAGAGAGTACTGCAAGGACCTGAATGCCTCAG
ACAACAACACCGAGTTCCTGAAAACTTTATTGAGCTCATGGAGAAAGTGACTCCAGACTC
CAAGCAGTGCAACAACCTTCCTTCTGCACAACCTGATCTTGGACACGGGCATCACGCAGCAG
CTGGTAGAGCGTGTGTGGAGGGACCAGGATCTCAACACGTACAGCCTACTGGCCGTGTTTCG
CTGCCACAGACGGTGGCATCACCCGAGTCTTCCCCAACAAAGGCAGCTGAGGACTGGACAGA
20 GAACCCTGAGCCCTTCAATGCCAGCTTCTACCGCCGCAGCCTGGATAACCACGGTTATGTC
TTCAAGCCCCCACACCAGGATGCCCTGTTAAGGCGCTGGAGCTGGAGAATGACACTGTGG
GCATCCTCGTCAGCACAGCTGTGGAGCTCAGCCTAGGCAGGCGCACACTGAGGCCAGCAGT
GGTGGGCGTCAAGCTGGACCTAGAGGCTTGGGCTGAGAAGTTCAAGGTGCTAGCCAGCAAC
CGTACCCACCAAGACCAGCCTCAGAAGTGCGGCCCAACAGCCACTGTGAGATGGACTGCG
25 AGGTTAACAATGAGGACTTACTCTGTGTCCTCATTGATGATGGAGGATTCTGGTGCTGTC
AAACCAGAACCATCAGTGGGACCAGGTGGGCAGGTTCTTCAGTGAGGTGGATGCCAACCTG
ATGCTGGCACTCTACAATAACTCCTTCTACACCGCAAGGAGTCCATGACTATCAGGCAG
CCTGTGCCCCCTCAGCCCCCTGGCAACCTGGGTGCTGCACCCCGGGGTGTCTTTGTGCCAC
CGTTGCAGATTTCCCTTAACCTGGCCTGGTGACCTCTGCTGCCGCCTGGTCCCTGTTCCAG
30 CAGCTTCTCTACGGCCTCATCTACCACAGCTGGTTCCAAGCAGACCCCGCGGAGGCCGAGG
GGAGCCCCGAGACGCGCGAGAGCAGCTGCGTCATGAAACAGACCCAGTACTACTTCGGCTC
GGTAAACGCCTCCTACAACGCCATCATCGACTGCGGAAACTGCTCCAGGCTGTTCCACGCG
CAGAGACTGACCAACACCAATCTTCTCTTTGTGGTGGCCGAGAAGCCGCTGTGCAGCCAGT
GCGAGGCTGGCCGGCTGCTGCAGAAGGAGACGCACTGCCAGCGGACGGCCCGGAGCAGTG
35 TGAGCTAGTGCAGAG

ATGGCGGTGTCGGCTCGGACCTGCGCGCCCTCTCGGCCCGGCCAGCGGACTGCGCGCC
CCTGGCCCCGGCTGCGGCCCCCACCCTGGCCCCGGCACCCGGCGCCCGACGTCCGGGCCCCC
GCGCCCGCTGTGGCTGCTGCTGCCGCTTCTACCGCTGCTCGCCGCCCCCGGCGCCTCTGCC
5 TACAGCTTCCCCCAGCAGCACACGATGCAGCACTGGGCCCCGGCGTCTGGAGCAGGAGGTGCG
ACGGCGTGATGCGGATTTTGGAGGCGTCCAGCAGCTCCGTGAGATTTACAAGGACAACCG
GAACCTGTTTCGAGGTACAGGAGAATGAGCCTCAGAAGTTGGTGGAGAAGGTGGCAGGGGAC
ATTGAGAGCCTTCTGGACAGGAAGGTGCAGGCCCTGAAGAGACTGGCTGATGCTGCAGAGA
ACTTCCAGAAAGCACACCGCTGGCAGGACAACATCAAGGAGGAAGACATCGTGTACTATGA
10 CGCCAAGGCTGACGCTGAGCTGGACGACCCTGAGAGTGAGGATGTGGAAAGGGGGTCTAAG
GCCAGCACCTTAAGGCTGGACTTCATCGAGGACCCAACTTCAAGAACAAGGTCAACTATT
CATACGCGGCTGTACAGATCCCTACGGACATCTACAAAGGCTCCACTGTCATCCTCAATGA
GCTCAACTGGACAGAGGCCCTGGAGAATGTGTTTCATGGAAAACCGCAGACAAGACCCCCACA
CTGCTGTGGCAGGTCTTCGGCAGCGCCACAGGAGTCACTCGCTACTACCCGGCCACCCCGT
15 GCGGAGCCCCCAAGAAGATCGACCTGTACGATGTCCGAAGGAGACCCTGGTATATCCAGGG
GGCCTCGTCACCCAAAGACATGGTTCATCATCGTGGATGTGAGTGGCAGTGTGAGCGGCCTG
ACCCTGAAGCTGATGAAGACATCTGTCTGCGAGATGCTGGACACGCTGTCTGATGATGACT
ATGTGAATGTGGCCTCGTTCAACGAGAAGGCACAGCCTGTGTTCATGCTTACACACCTGGT
GCAGGCCAATGTGCGCAACAAGAAGGTGTTCAAGGAAGCTGTGCAGGGCATGGTGGCCAAG
20 GGCACCACAGGCTACAAGGCCGGCTTTGAGTATGCCTTTGACCAGCTGCAGAACTCCAACA
TCACTCGGGCCAACTGCAACAAGATGATCATGATGTTACGGATGGTGGTGAGGACCGCGT
GCAGGACGTCTTTGAGAAGTACAATTGGCCAAACCGGACGGTGCGCGTGTTTACTTTCTCC
GTGGGGCAGCATAACTATGACGTACACCCGCTGCAGTGGATGGCCTGTGCCAACAAGGCT
ACTATTTTGAGATCCCTTCCATCGGAGCCATCCGCATCAACACACAGGAATATCTAGATGT
25 GTTGGGCAGGCCCATGGTGTGCTGGCAGGCAAGGAGGCCAAGCAGTTTCAGTGGACCAACGTG
TATGAGGATGCAC'TGGGACTGGGGTTGGTGGTAACAGGGACCCTCCCTGTTTTCAACCTGA
CACAGGATGGCCCTGGGGAAAAGAAGAACAGCTGATCCTGGGCGTGATGGGCATTGACGT
GGCTCTGAATGACATCAAGAGGCTGACCCCCAACTACACGCTTGGAGCCAACGGCTATGTG
TTTGCCATTGACCTGAACGGCTACGTGTTGCTGCACCCCAATCTCAAGCCCCAGACCACCA
30 ACTTCCGGGAGCCTGTGACTCTGGACTTCCTGGATGCGGAGCTAGAGGATGAGAACAAGGA
AGAGATCCGTCGGAGCATGATTTGATGGCAACAAGGGCCACAAGCAGATCAGAACGTTGGTC
AAGTCCCTGGATGAGAGGTACATAGATGAGGTGACACGGAACTACACCTGGGTGCCTATAA
GGAGCACTAACTACAGCCTGGGGCTGGTGCTCCACCCCTACAGCACCTTCTACCTCCAAGC
CAATCTCAGTGACCAGATCCTGCAGGTCAAGTATTTTGAGTTCCTGCTCCCCAGCAGCTTT
35 GAGTCTGAAGGACACGTTTTTCATTGCTCCAGAGAGTACTGCAAGGACCTGAATGCCTCAG
ACAACAACACCGAGTTCTTGAAAACTTTATTGAGCTCATGGAGAAAGTGACTCCAGACTC
CAAGCAGTGCAACAACCTTCTTCTGCACAACCTGATCTTGGACACGGGCATCACGCAGCAG
CTGGTAGAGCGTGTGTGGAGGGACCAGGATCTCAACACGTACAGCCTACTGGCCGTGTTCC

CTGCCACAGACGGTGGCATCACCCGAGTCTTCCCCAACAAAGGCAGCTGAGGACTGGACAGA
 GAACCCTGAGCCCTTCAATGCCAGCTTCTACCGCCGAGCCTGGATAAACCACGGTTATGTC
 TTCAAGCCCCACACCAGGATGCCCTGTTAAGGCCGCTGGAGCTGGAGAATGACACTGTGG
 GCATCCTCGTCAGCACAGCTGTGGAGCTCAGCCTAGGCAGGCGCACACTGAGGCCAGCAGT
 5 GGTGGGCGTCAAGCTGGACCTAGAGGCTTGGGCTGAGAAGTTCAAGGTGCTAGCCAGCAAC
 CGTACCCACCAAGACCAGCCTCAGAAGTGCGGCCCCAACAGCCACTGTGAGATGGACTGCG
 AGGTTAACAATGAGGACTTACTCTGTGTCTCATTTGATGATGGAGGATTCCTGGTGCTGTC
 AAACCAGAACCATCAGTGGGACCAGGTGGGCAGGTTCTTCAGTGAGGTGGATGCCAACCTG
 ATGCTGGCACTCTACAATAACTCCTTCTACACCCGCAAGGAGTCCATGACTATCAGGCAG
 10 CCTGTGCCCCCTCAGCCCCCTGGCAACCTGGGTGCTGCACCCCGGGGTGTCTTTGTGCCCCAC
 CGTTGCAGATTTCTTAACCTGGCCTGGTGGACCTCTGCTGCCGCTGGTCCCTGTTCCAG
 CAGCTTCTCTACGGCCTCATCTACCACAGCTGGTTCCAAGCAGACCCCGCGGAGGCCGAGG
 GGAGCCCCGAGACGCGCGAGAGCAGCTGCGTCATGAAACAGACCCAGTACTACTTCGGCTC
 GGTAAACGCCTCCTACAACGCCATCATCGACTGCGGAAACTGCTCCAGGCTGTTCCACGCG
 15 CAGAGACTGACCAACACCAATCTTCTCTTTGTGGTGGCCGAGAAGCCGCTGTGCAGCCAGT
 GCGAGGCTGGCCGCTGCTGCAGAAGGAGACGCACTGCCAGCGGACGGCCCGGAGCAGTG
 TGAGCTAGTGCAGAGACCGGATACCGGAGAGGCCCGCACATCTGCTTCGACTACAACGCG
 ACAGAAGATACCTCAGACTGTGGCCGCGGGGCC

20 **4 - human amino acid sequence**

MAVPARTCGASRPGPARTARPWPGCGPHPGPGTRRPTSGPPRPLWLLLPLLLAAPGASA
 YSFPQQHTMQHWARRLEQEVDGVMRIFGGVQQLREIYKDNRLFEVQENEPQKLVEKVAGD
 IESLLDRKVQALKRLADAAENFQKAHRWQDNIKEEDIVYYDAKADAELDDPESEDVERGSK
 ASTLRLDFIEDPNFKNKNVYSYAAVQIPTDIYKGSTVILNELNWTEALENVFMENRRQDPT
 25 LLWQVFGSATGVTRYYPATPWRAKKIDLYDVRRRPWYIQGASSPKDMVIIVDVSGSVSGL
 TLKLMKTSVCEMLDTLSDDDYVNVA SFNEKAQPVSCFTHLVQANVRNKKVFKEAVQGMVAK
 GTTGKAGFEYAFDQLQNSNITRANCNMIMMFTDGGEDRVQDVFEKYNWPNRTRVFTFS
 VGQHNYDVTPLQWMACANKGYFFEIP SIGAIRINTQEYLDVLGRPMVLAGEAKQVQWTVN
 YEDALGLGLVVTGTLPVFNLTQDGPGEKKQLILGVMGIDVALNDIKRLTPNYTLGANGYV
 30 FAIDLNGYVLLHPNLKPQTTNFREPVTLDFLDAELEDENKEEIRRS MIDGNKGHKQIRTLV
 KSLDERYIDEVTRNYTWVPIRSTNYSGLVLPPYSTFYLQANLSDQILQVKYFEFLPSSF
 ESEGHVFIAPREYCKDLNASDNNTFLKNFIELMEKVTPDSKQCNNFLLHNLILDTGITQQ
 LVERVWRDQDLNTYSLLAVFAATDGGITRVFPNKAEDWTENPEPFNASFYRRSLDNHGYV
 FKPPHQDALLRPLELENDTVGILVSTAVELSLGRRTLRPAVVGKLDLEAWAEKFKVLASN
 35 RTHQDQPQKCGPNSHCEMDCEVNNEDLLCVLIDDGGFLVLSNQNHQWDQVGRFFSEVDANL
 MLALYNNSFYTRKESYDYQAACAPQPPGNLGAAPRGVFPVPTVADFLNLAWWTSAAAWSLFQ
 QLLYGLIYHSWFQADPAEAGSPETRESSCVMKQTQYYFGSVNASYNAIIDCGNCSRLFHA
 QRLTNTNLLFVVAEKPLCSQCEAGR

bioRxiv preprint doi: <https://doi.org/10.1101/099266>; this version posted September 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

5 - human amino acid sequence

MAVPARTCGASRPGPARTARPWPGCGPHPGPGTRRPTSGPPRPLWLLLPLLPLLAAPGASA
 YSFPQQHTMQHWARRLEQEVDGVMRIFGGVQQQLREIYKDNRNLFVQENEPQKLVEKVAGD
 IESLLDRKVQALKRLADAAENFQKAHRWQDNIKEEDIVYYDAKADAELDDPESEDVERGSK
 5 ASTLRDLDFIEDPNFKNKNVNSYAAVQIPTDIYKGSTVILNELNWTEALENVFMENRRQDPT
 LLWQVFGSATGVTRYYPATPWRAPKKIDLYDVRRRPWYIQGASSPKDMVIIVDVSGSVSGL
 TLKLMKTSVCEMLDTLSDDDYVNVASFNEKAQPVSCFTHLVQANVRNKKVFKEAVQGMVAK
 GTTGYSKAGFEYAFDQLQNSNITRANCNKMMFTDGGEDRVQDVFEKYNWPNRTRVRVFTFS
 VGQHNVDVTPLQWMACANKGYFFEIPSIGAIRINTQEYLDVLGRPMVLAGEAKQVQWNTV
 10 YEDALGLGLVVTGTLPVFNLTQDGPGEKKNQLILGVMGIDVALNDIKRLTPNYTLGANGYV
 FAIDLNGYVLLHPNLKPQTTNFREPVTLDFLDAELEDENKEEIRRSIDGNKGGHKQIRTLV
 KSLDERYIDEVTRNYTWPIRSTNYSGLVLPYSTFYLQANLSDQILQVKYFEFLLPSSF
 ESEGHVFIAPREYCKDLNASDNNTEFLKNFIELMEKVT PDSKQCNNFLLHNLILDTGITQQ
 LVERVWRDQDLNTYSLLAVFAATDGGITRVFPNKAEDWTENPEPFNASFYRRSLDNHGYV
 15 FKPPHQDALLRPLELENDTVGILVSTAVELSLGRRTLRAVVGKLDLEAWAEKFKVLASN
 RTHQDQPKCGPNSHCEMDCEVNNEDELLCVLIDDGGFLVLSNQNHQWDQVGRFFSEVDANL
 MLALYNNSFYTRKESYDYQAACAPQPPGNLGAAPRGVFPVPTVADFLNLAWWTSAAAWSLFQ
 QLLYGLIYHSWFQADPAEAEGSPETRESSCVMKQTQYYFGSVNASYNAIIDCGNCSRLFHA
 QRLTNTNLLFVVAEKPLCSQCEAGRLLQKETHCPADGPEQCELVQ

20

6 - human amino acid sequence

MAVPARTCGASRPGPARTARPWPGCGPHPGPGTRRPTSGPPRPLWLLLPLLPLLAAPGASA
 YSFPQQHTMQHWARRLEQEVDGVMRIFGGVQQQLREIYKDNRNLFVQENEPQKLVEKVAGD
 IESLLDRKVQALKRLADAAENFQKAHRWQDNIKEEDIVYYDAKADAELDDPESEDVERGSK
 25 ASTLRDLDFIEDPNFKNKNVNSYAAVQIPTDIYKGSTVILNELNWTEALENVFMENRRQDPT
 LLWQVFGSATGVTRYYPATPWRAPKKIDLYDVRRRPWYIQGASSPKDMVIIVDVSGSVSGL
 TLKLMKTSVCEMLDTLSDDDYVNVASFNEKAQPVSCFTHLVQANVRNKKVFKEAVQGMVAK
 GTTGYSKAGFEYAFDQLQNSNITRANCNKMMFTDGGEDRVQDVFEKYNWPNRTRVRVFTFS
 VGQHNVDVTPLQWMACANKGYFFEIPSIGAIRINTQEYLDVLGRPMVLAGEAKQVQWNTV
 30 YEDALGLGLVVTGTLPVFNLTQDGPGEKKNQLILGVMGIDVALNDIKRLTPNYTLGANGYV
 FAIDLNGYVLLHPNLKPQTTNFREPVTLDFLDAELEDENKEEIRRSIDGNKGGHKQIRTLV
 KSLDERYIDEVTRNYTWPIRSTNYSGLVLPYSTFYLQANLSDQILQVKYFEFLLPSSF
 ESEGHVFIAPREYCKDLNASDNNTEFLKNFIELMEKVT PDSKQCNNFLLHNLILDTGITQQ
 LVERVWRDQDLNTYSLLAVFAATDGGITRVFPNKAEDWTENPEPFNASFYRRSLDNHGYV
 35 FKPPHQDALLRPLELENDTVGILVSTAVELSLGRRTLRAVVGKLDLEAWAEKFKVLASN
 RTHQDQPKCGPNSHCEMDCEVNNEDELLCVLIDDGGFLVLSNQNHQWDQVGRFFSEVDANL
 MLALYNNSFYTRKESYDYQAACAPQPPGNLGAAPRGVFPVPTVADFLNLAWWTSAAAWSLFQ
 QLLYGLIYHSWFQADPAEAEGSPETRESSCVMKQTQYYFGSVNASYNAIIDCGNCSRLFHA

093750:05520560

QRLTNTNLLFVVAEKPLCSQCEAGRLLQKETHCPADGPEQCELVQRPRYRRGPHICFDYNA
TEDTSDCGRGA

7 - human nucleotide sequence

5 ATGGCCGGGCGGGCTCGCCGCGCCGCGCTCCCGGGGGCCTCGGCGCTTCTCGCTGCCG
CGCTTCTCTACGCCGCGCTGGGGGACGTGGTGCGCTCGGAGCAGCAGATACCGCTCTCCGT
GGTGAAGCTCTGGGCCTCGGCTTTTGGTGGGGAGATAAAATCCATTGCTGCTAAGTACTCC
GGTTCCACGCTTCTGCAAAAGAAATACAAAGAGTATGAGAAAGACGTTGCCATAGAAGAAA
TTGATGGCCTCCAACCTGGTAAAGAAGCTGGCAAAGAACATGGAAGAGATGTTTCACAAGAA
10 GTCTGAGGCCGTCAGGCGTCTGGTGAGGCTGCAGAAGAAGCACACCTGAAACATGAATTT
GATGCAGACTTACAGTATGAATACTTCAATGCTGTGCTGATAAATGAAAGGGACAAAGACG
GGAATTTTTTGGAGCTGGGAAAGGAATTCATCTTAGCCCCAAATGACCATTTTAATAATTT
GCCTGTGAACATCAGTCTAAGTGACGTCCAAGTACCAACGAACATGTACAACAAAGACCCT
GCAATTGTCAATGGGGTTTATTGGTCTGAATCTCTAAACAAAGTTTTTGTAGATAACTTTG
15 ACCGTGACCCATCTCTCATATGGCAGTACTTTGGAAGTGCAAAGGGCTTTTTTAGGCAGTA
TCCGGGGATTAAATGGGAACCAGATGAGAATGGAGTCATTGCCTTCGACTGCAGGAACCGA
AAATGGTACATCCAGGCAGCAACTTCTCCGAAAGACGTGGTCATTTTAGTTGACGTCAGTG
GCAGCATGAAAGGACTCCGTCTGACTATCGCGAAGCAAACAGTCTCATCCATTTTGGATAC
ACTTGGGGATGATGACTTCTTCAACATAATTGCTTATAATGAGGAGCTTCACTATGTGGAA
20 CCTTGCTGAATGGAACTTTGGTGCAAGCCGACAGGACAAACAAAGAGCACTTCAGGGAGC
ATCTGGACAAACTTTTCGCCAAAGGAATTGGAATGTTGGATATAGCTCTGAATGAGGCCTT
CAACATTCTGAGTGATTTCAACCACACGGGACAAGGAAGTATCTGCAGTCAGGCCATCATG
CTCATAACTGATGGGGCGGTGGACACCTATGATACAATCTTTGCAAAATACAATTGGCCAG
ATCGAAAGGTTTCGCATCTTCACATACCTCATTTGGACGAGAGGCTGCGTTTGCAGACAATCT
25 AAAGTGGATGGCCTGTGCCAACAAAGGATTTTTTACCCAGATCTCCACCTTGGCTGATGTG
CAGGAGAATGTCATGGAATACCTTCACGTGCTTAGCCGGCCCAAAGTCATCGACCAGGAGC
ATGATGTGGTGTGGACCGAAGCTTACATTGACAGCACTCTGACTGATGATCAGGGCCCCGT
CCTGATGACCCTGTAGCCATGCCTGTGTTTAGTAAGCAGAACGAAACCAGATCGAAGGGC
ATTCTTCTGGGAGTGGTTGGCACAGATGTCCCAGTGAAAGAACTTCTGAAGACCATCCCCA
30 AATACAAGTTAGGGATTACGGTTATGCCTTTGCAATCACAAATAATGGRTATATCCTGAC
GCATCCGGAACTCAGGCTGCTGTACGAAGAAGGAAAAAGCGAAGGAAACCTAACTATAGT
AGCGTTGACCTCTCTGAGGTGGAGTGGGAAGACCGAGATGACGTGTTGAGAAATGCTATGG
TGAATCGAAAGACGGGGAAGTTTTCCATGGAGGTGAAGAAGACAGTGGACAAAGGGAAACG
GGTTTTGGTGATGACAAATGACTACTATTATACAGACATCAAGGGTACTCCTTTCAGTTTA
35 GGTGTGGCGCTTTCCAGAGGTCATGGGAAATATTTCTTCCGAGGGAATGTAACCATCGAAG
AAGGCCTGCATGACTTAGAACATCCCGATGTGTCTTGGCAGATGAATGGTCCTACTGCAA
CACTGACCTACACCCTGAGCACCGCCATCTGTCTCAGTTAGAAGCGATTAAGCTCTACCTA
AAAGGCAAAGAACCTCTGCTCCAGTGTGATAAAGAATTGATCCAAGAAGTCCTTTTTTGACG

009760"05526660

CGGTGGTGAGTGCCCCATTGAAGCGTATTGGACCAGCCTGGCCCTCAACAAATCTGAAAA
 TTCTGACAAGGGCGTGGAGGTTGCCCTTCCTCGGCACCTCGCACGGGCCTCTCCAGAATCAAC
 CTGTTTGTGCGGGCTGAGCAGCTCACCAATCAGGACTTCCTGAAAGCTGGCGACAAGGAGA
 ACATTTTTTAACGCAGACCATTTCCCTCTCTGGTACCGAAGAGCCGCTGAGCAGATTCCAGG
 5 GAGCTTCGTCTACTCGATCCCATTTCAGCACTGGACCAGTCAATAAAAGCAATGTGGTGACA
 GCAAGTACATCCATCCAGCTCCTGGATGAACGGAAATCTCCTGTGGTGGCAGCTGTAGGCA
 TTCAGATGAACTTGAATTTTTTCCAAAGGAAGTTCTGGACTGCCAGCAGACAGTGTGCTTC
 CCTGGATGGCAAATGCTCCATCAGCTGTGATGATGAGACTGTGAATTGTTACCTCATAGAC
 AATAATGGATTTATTTTGGTGTCTGAAGACTACACACAGACTGGAGACTTTTTTGGTGAGA
 10 TCGAGGGAGCTGTGATGAACAAATTGCTAACAATGGGCTCCTTTAAAAGAATTACCCTTTA
 TGA CTACCAAGCCATGTGTAGAGCCAACAAGGAAAGCAGCGATGGCGCCCATGGCCTCCTG
 GATCCTTATAATGCCTTCCTCTCTGCAGTAAAATGGATCATGACAGAACTTGTCTTGTTC
 TGGTGGAATTTAACCTCTGCAGTTGGTGGCACTCCGATATGACAGCTAAAGCCCAGAAATT
 GAAACAGACCCTGGAGCCTTGTGATACTGAATATCCAGCATTTCGTCTCTGAGCGCACCATC
 15 AAGGAGACTACAGGGAATATTGCTTGTGAAGACTGCTCCAAGTCCTTTGTTCATCCAGCAA
 TCCAAGCAGCAACCTGTTTCATGGTGGTGGTGGACAGCAGCTGCCTCTGTGAATCTGTGGC
 CCCCATC

8 - human nucleotide sequence

20 ATGGCCGGGCCGGGCTCGCCGCGCCGCGCTCCCGGGGGCCTCGGCGCTTCTCGCTGCCG
 CGCTTCTCTACGCCGCGTGGGGGACGTGGTGCCTCGGAGCAGCAGATACCGCTCTCCGT
 GGTGAAGCTCTGGGCCTCGGCTTTTGGTGGGGAGATAAAATCCATTGCTGCTAAGTACTCC
 GGTTCCTCAGCTTCTGCAAAGAAATACAAAGAGTATGAGAAAGACGTTGCCATAGAAGAAA
 TTGATGGCCTCCAACCTGGTAAAGAAGCTGGCAAAGAACATGGAAGAGATGTTTCACAAGAA
 25 GTCTGAGGCCGTCAGGCGTCTGGTGGAGGCTGCAGAAGAAGCACACCTGAAACATGAATTT
 GATGCAGACTTACAGTATGAATACTTCAATGCTGTGCTGATAAATGAAAGGGACAAAGACG
 GGAATTTTTTGGAGCTGGGAAAGGAATTCATCTTAGCCCCAAATGACCATTTTAATAATTT
 GCCTGTGAACATCAGTCTAAGTGACGTCCAAGTACCAACGAACATGTACAACAAAGACCCT
 GCAATTGTCAATGGGGTTTATTGGTCTGAATCTCTAAACAAAGTTTTTGTAGATAACTTTG
 30 ACCGTGACCCATCTCTCATATGGCAGTACTTTGGAAGTGCAAAGGGCTTTTTTAGGCAGTA
 TCCGGGGATTAAATGGGAACCAGATGAGAATGGAGTCATTGCCTTCGACTGCAGGAACCGA
 AAATGGTACATCCAGGCAGCAACTTCTCCGAAAGACGTGGTCATTTTAGTTGACGTCAGTG
 GCAGCATGAAAGGACTCCGTCTGACTATCGCGAAGCAAACAGTCTCATCCATTTTGGATAC
 ACTTGGGGATGATGACTTCTTCAACATAATTGCTTATAATGAGGAGCTTCACTATGTGGAA
 35 CCTTGCCTGAATGGAACTTTGGTGCAAGCCGACAGGACAAACAAAGAGCACTTCAGGGAGC
 ATCTGGACAAACTTTTCGCCAAAGGAATTGGAATGTTGGATATAGCTCTGAATGAGGCCTT
 CAACATTCTGAGTGATTTCAACCACACGGGACAAGGAAGTATCTGCAGTCAGGCCATCATG
 CTCATAACTGATGGGGCGGTGGACACCTATGATACAATCTTTGCAAATACAATTGGCCAG

09397550:091599

ATCGAAAGGTTTCGCATCTTCACATACCTCATTGGACGAGAGGCTGCGTTTGCAGACAATCT
 AAAGTGGATGGCCTGTGCCAACAAAGGATTTTTTACCCAGATCTCCACCTTGGCTGATGTG
 CAGGAGAATGTCATGGAATACCTTCACGTGCTTAGCCGGCCCAAAGTCATCGACCAGGAGC
 ATGATGTGGTGTGGACCGAAGCTTACATTGACAGCACTCTGACTGATGATCAGGGCCCCGT
 5 CCTGATGACCACTGTAGCCATGCCTGTGTTTAGTAAGCAGAACGAAACCAGATCGAAGGGC
 ATTCTTCTGGGAGTGGTTGGCACAGATGTCCCAGTGAAAGAACTTCTGAAGACCATCCCCA
 AATACAAGTTAGGGATTACGGTTATGCCTTTGCAATCACAAATAATGGRTATATCCTGAC
 GCATCCGGAACCTCAGGCTGCTGTACGAAGAAGGAAAAAGCGAAGGAAACCTAACTATAGT
 AGCGTTGACCTCTCTGAGGTGGAGTGGGAAGACCGAGATGACGTGTTGAGAAATGCTATGG
 10 TGAATCGAAAGACGGGGAAGTTTTCCATGGAGGTGAAGAAGACAGTGGACAAAGGGAAACG
 GGTTTTGGTGATGACAAATGACTACTATTATACAGACATCAAGGGTACTCCTTTTCAGTTTA
 GGTGTGGCGCTTTCCAGAGGTCATGGGAAATATTTCTTCCGAGGGAATGTAACCATCGAAG
 AAGGCCTGCATGACTTAGAACATCCCGATGTGTCTTGGCAGATGAATGGTCTTACTGCAA
 CACTGACCTACACCCTGAGCACCGCCATCTGTCTCAGTTAGAAGCGATTAAGCTCTACCTA
 15 AAAGGCAAAGAACCTCTGCTCCAGTGTGATAAAGAATTGATCCAAGAAGTCCTTTTTGACG
 CGGTGGTGAGTGCCCCATTGAAGCGTATTGGACCAGCCTGGCCCTCAACAAATCTGAAAA
 TTCTGACAAGGGCGTGGAGGTTGCCTTCCCTCGGCACCTCGCACGGGCCTCTCCAGAATCAAC
 CTGTTTGTGCGGGCTGAGCAGCTCACCAATCAGGACTTCCTGAAAGCTGGCGACAAGGAGA
 ACATTTTTTAACGCAGACCATTTCCCTCTCTGGTACCGAAGAGCCGCTGAGCAGATTCCAGG
 20 GAGCTTCGTCTACTCGATCCCATTCAGCACTGGACCAGTCAATAAAAGCAATGTGGTGACA
 GCAAGTACATCCATCCAGCTCCTGGATGAACGGAAATCTCCTGTGGTGGCAGCTGTAGGCA
 TTCAGATGAAACTTGAATTTTTCCAAAGGAAGTTCTGGACTGCCAGCAGACAGTGTGCTTC
 CCTGGATGGCAAATGCTCCATCAGCTGTGATGATGAGACTGTGAATTGTTACCTCATAGAC
 AATAATGGATTTATTTTGGTGTCTGAAGACTACACACAGACTGGAGACTTTTTTGGTGAGA
 25 TCGAGGGAGCTGTGATGAACAAATTGCTAACAATGGGCTCCTTTAAAGAATTACCCTTTA
 TGACTIONACCAAGCCATGTGTAGAGCCAACAAGGAAAGCAGCGATGGCGCCCATGGCCTCCTG
 GATCCTTATAATGCCTTCCTCTCTGCAGTAAAATGGATCATGACAGAACTTGTCTTGTTC
 TGGTGGAATTTAACCTCTGCAGTTGGTGGCACTCCGATATGACAGCTAAAGCCCAGAAATT
 GAAACAGACCCTGGAGCCTTGTGATACTGAATATCCAGCATTTCGTCTCTGAGCGCACCATC
 30 AAGGAGACTACAGGGAATATTGCTTGTGAAGACTGCTCCAAGTCCTTTGTTCATCCAGCAA
 TCCCAAGCAGCAACCTGTTTCATGGTGGTGGTGGACAGCAGCTGCCTCTGTGAATCTGTGGC
 CCCCATCACCATGGCACCCATTGAAATCAGGTATAATGAATCCCTTAAGTGTGAACGTCTA
 AAG

35 **9 - human nucleotide sequence**

ATGGCCGGGCGGGGCTCGCCGCGCCGCGCTCCCGGGGGCCTCGGCGCTTCTCGCTGCCG
 CGCTTCTCTACGCCGCGCTGGGGGACGTGGTGGCTCGGAGCAGCAGATACCGCTCTCCGT
 GGTGAAGCTCTGGGCCTCGGCTTTTGGTGGGGAGATAAAATCCATTGCTGCTAAGTACTCC

GGTCCAGCTTCTGCAAAAGAAATACAAAGAGTATGAGAAAGACGTTGCCATAGAAGAA
TTGATGGCCTCCAACCTGGTAAAGAAGCTGGCAAAGAACATGGAAGAGATGTTTCACAAGAA
GTCTGAGGCCGTGAGGCGTCTGGTGGAGGCTGCAGAAGAAGCACACCTGAAACATGAATTT
GATGCAGACTTACAGTATGAATACTTCAATGCTGTGCTGATAAATGAAAGGGACAAAGACG
5 GGAATTTTTTTGGAGCTGGGAAAGGAATTCATCTTAGCCCCAAATGACCATTTTAATAATTT
GCCTGTGAACATCAGTCTAAGTGACGTCCAAGTACCAACGAACATGTACAACAAAGACCCCT
GCAATTGTCAATGGGGTTTATTGGTCTGAATCTCTAAACAAAGTTTTTTGTAGATAACTTTG
ACCGTGACCCATCTCTCATATGGCAGTACTTTGGAAGTGCAAAGGGCTTTTTTAGGCAGTA
TCCGGGGATTAAATGGGAACCAGATGAGAATGGAGTCATTGCCTTCGACTGCAGGAACCGA
10 AAATGGTACATCCAGGCAGCAACTTCTCCGAAAGACGTGGTCATTTTAGTTGACGTCAGTG
GCAGCATGAAAGGACTCCGTCTGACTATCGCGAAGCAAACAGTCTCATCCATTTTGGATAC
ACTTGGGGATGATGACTTCTTCAACATAATTGCTTATAATGAGGAGCTTCACTATGTGGAA
CCTTGCCCTGAATGGAACCTTTGGTGCAGCCGACAGGACAAACAAAGAGCACTTCAGGGAGC
ATCTGGACAAACTTTTTCGCCAAAGGAATTGGAATGTTGGATATAGCTCTGAATGAGGCCTT
15 CAACATTCTGAGTGATTTCAACCACACGGGACAAGGAAGTATCTGCAGTCAGGCCATCATG
CTCATAACTGATGGGGCGGTGGACACCTATGATACAATCTTTGCAAAATACAATTGGCCAG
ATCGAAAGGTTGCGATCTTCACATACCTCATTTGGACGAGAGGCTGCGTTTGACAGACAATCT
AAAGTGGATGGCCTGTGCCAACAAAGGATTTTTTACCCAGATCTCCACCTTGGCTGATGTG
CAGGAGAATGTCATGGAATACCTTCACGTGCTTAGCCGGCCCAAAGTCATCGACCAGGAGC
20 ATGATGTGGTGTGGACCGAAGCTTACATTGACAGCACTCTGACTGATGATCAGGGCCCCGT
CCTGATGACCACTGTAGCCATGCCTGTGTTTAGTAAGCAGAACGAAACCAGATCGAAGGGC
ATTCTTCTGGGAGTGGTTGGCACAGATGTCCAGTGAAAGAACTTCTGAAGACCATCCCCA
AATACAAGTTAGGGATTACGGTTATGCCTTTGCAATCACAAATAATGGRTATATCCTGAC
GCATCCGGAACCTCAGGCTGCTGTACGAAGAAGGAAAAAAGCGAAGGAAACCTAACTATAGT
25 AGCGTTGACCTCTCTGAGGTGGAGTGGGAAGACCGAGATGACGTGTTGAGAAATGCTATGG
TGAATCGAAAGACGGGGAAGTTTTCCATGGAGGTGAAGAAGACAGTGGAACAAAGGGAAACG
GGTTTTGGTGATGACAAATGACTACTATTATACAGACATCAAGGGTACTCCTTTTCAGTTTA
GGTGTGGCGCTTTCCAGAGGTCATGGGAAATATTTCTTCCGAGGGAATGTAACCATCGAAG
AAGGCCTGCATGACTTAGAACATCCCGATGTGTCTTGGCAGATGAATGGTCCTACTGCAA
30 CACTGACCTACACCCTGAGCACCGCCATCTGTCTCAGTTAGAAGCGATTAAGCTCTACCTA
AAAGGCAAAGAACCTCTGCTCCAGTGTGATAAAGAATTGATCCAAGAAGTCCTTTTTGACG
CGGTGGTGAGTGCCCCATTGAAGCGTATTGGACCAGCCTGGCCCTCAACAAATCTGAAAA
TTCTGACAAGGGCGTGGAGGTTGCCTTCTCGGCACTCGCACGGGCCTCTCCAGAATCAAC
CTGTTTGTGCGGGCTGAGCAGCTCACCAATCAGGACTTCTGAAAGCTGGCGACAAGGAGA
35 ACATTTTTTAACGCAGACCATTTCCTCTCTGGTACCGAAGAGCCGCTGAGCAGATTCCAGG
GAGCTTCGTCTACTCGATCCCATTCAGCACTGGACCAGTCAATAAAAGCAATGTGGTGACA
GCAAGTACATCCATCCAGCTCCTGGATGAACGGAAATCTCCTGTGGTGGCAGCTGTAGGCA
TTCAGATGAACTTGAATTTTTTCAAAGGAAGTTCTGGACTGCCAGCAGACAGTGTGCTTC

CCTGGATGGCAAATGCTCCATCAGCTGTGATGATGAGACTGTGAATTGTTACCTCATAGAC
 AATAATGGATTTATTTTGGTGTCTGAAGACTACACACAGACTGGAGACTTTTTTGGTGAGA
 TCGAGGGAGCTGTGATGAACAAATTGCTAACAATGGGCTCCTTTAAAAGAATTACCCTTTA
 TGACTACCAAGCCATGTGTAGAGCCAACAAGGAAAGCAGCGATGGCGCCCATGGCCTCCTG
 5 GATCCTTATAATGCCTTCCTCTCTGCAGTAAAATGGATCATGACAGAACTTGTCTTGTTC
 TGGTGGAATTTAACCTCTGCAGTTGGTGGCACTCCGATATGACAGCTAAAGCCCAGAAATT
 GAAACAGACCCTGGAGCCTTGTGATACTGAATATCCAGCATTCGTCTCTGAGCGCACCATC
 AAGGAGACTACAGGGAATATTGCTTGTGAAGACTGCTCCAAGTCCTTTGTCATCCAGCAAA
 TCCCAAGCAGCAACCTGTTTCATGGTGGTGGTGGACAGCAGCTGCCTCTGTGAATCTGTGGC
 10 CCCCATCACCATGGCACCCATTGAAATCAGGTATAATGAATCCCTTAAGTGTGAACGTCTA
 AAGGCCCAGAAGATCAGAAGGGCCCAGAAGATCAGAAGGCGCCAGAATCTTGTTCATGGCT
 TCCATCCTGAGGAGAATGCAAGGGAGTGTGGGGGTGCGCCG

10 - human amino acid sequence

15 MAGPGSPRRASRGASALLAAALLYAALGDVVRSEQQIPLSVVKLWASAFGGEIKSIAAKYS
 GSQLLQKKYKEYEKDVAIEEIDGLQLVKKLAKNMEEMFHHKSEAVRRLVEAAEEAHLKHEF
 DADLQYEFNAVLINERDKDGNFLELGKEFILAPNDHFNNLPVNISLSDVQVPTNMYNKDP
 AIVNGVYWSESLNKVFVDNFDRLPSLIWQYFGSAKGFFRQYPGIKWEPDENGVIADFDCRNR
 KWIQAATSPKDVVILVDVSGSMKGLRLTIKQTVSSILDTLGDDDDFNIIAYNEELHYVE
 20 PCLNGTLVQADRTNKEHFREHLDKLFAKGIGMLDIALNEAFNILSDFNHTGQGSICSQAIM
 LITDGAVDTYDTIFAKYNWPDRKVRIFTYLGREAFAFADNLKWMACANKGFFTQISTLADV
 QENVMEYLHVLSPKVIDQEHDDVVWTEAYIDSTLTDDQGPVLMTTVAMPVFSKQNETRSKG
 ILLGVVGTDPVKELLKTIPKYKLGIGHYAFAITNNGYILTHPELRLLYEKGKRRKPNYS
 SVDLSEVEWEDRDDVLRNAMVNRKTGKFSMEVKKTVDKGKRVLVMNDYYYTDIKGTPFSL
 25 GVALSRGHGKYFFRGNVTIEEGLHLEHPDVSLADEWSYCNTDLHPEHRHLSQLEAIKLYL
 KGKEPLLQCDKELIQEVLFDAVVSAPIEAYWTSALNKSSENSDKGVEVAFLGTRTGLSRIN
 LFVGAEQLTNQDFLKAGDKENIFNADHFPLWYRRAAEQIPGSFVYSIPFSTGPVNKSNVVT
 ASTSIQLLDERKSPVVAAVGIQMKLEFFQRKFWTASRQCASLDGKCSISCDDDETVCYLID
 NNGFILVSEDTQTGDDFFGEIEGAVMNKLLTMGSFKRITLYDYQAMCRANKESSDGAHGLL
 30 DPYNAFLSAVKWIMTELVLFLVEFNLCSWWHSMDTAKAQKQKQTLPCDTEYPAFVSERTI
 KETTGNIACEDCSKSFVIQQIPSSNLFMVVDSSCLCESVAPI

11 - human amino acid sequence

MAGPGSPRRASRGASALLAAALLYAALGDVVRSEQQIPLSVVKLWASAFGGEIKSIAAKYS
 35 GSQLLQKKYKEYEKDVAIEEIDGLQLVKKLAKNMEEMFHHKSEAVRRLVEAAEEAHLKHEF
 DADLQYEFNAVLINERDKDGNFLELGKEFILAPNDHFNNLPVNISLSDVQVPTNMYNKDP
 AIVNGVYWSESLNKVFVDNFDRLPSLIWQYFGSAKGFFRQYPGIKWEPDENGVIADFDCRNR
 KWIQAATSPKDVVILVDVSGSMKGLRLTIKQTVSSILDTLGDDDDFNIIAYNEELHYVE

002750-00460

PCLNGTLVQADRTNKEHFREHLDKLFAKGIGMLDIALNEAFNILSDFNHTGQGSICSQAIM
 LITDGAVDTYDTIFAKYNWPDRKVRIFTYILIGREAAAFADNLKWMACANKGFFTQISTLADV
 QENVMEYLHVLSRPKVIDQEHDVVWTEAYIDSTLTDDQGPVLMTTVAMPVFSKQNETRSKG
 ILLGVVGTDPVPKELLKTIPKYKLGIGHYAFAITNNGYILTHPELRLLYEKGKRRKPNYS
 5 SVDLSEVEWEDRDDVLRNAMVNRKTGKFSMEVKKTVDKGKRVLMVTNDYYYTDIKGTPFSL
 GVALSRGHGKYFFRGNVTIEEGLHDLEHPDVSLADEWSYCNTDLHPEHRHLSQLEAIKLYL
 KGKEPLLQCDKELIQEVLFDAVVSAPIEAYWTSALNKSSENSDKGVEVAFLGTRTGLSRIN
 LFVGAEQLTNQDFLKAGDKENIFNADHFPLWYRRAAEQIPGSFVYSIPFSTGPVNKSNVVT
 ASTSIQLLDERKSPVVAAVGIQMKLEFFQRKFWTASRQCASLDGKCSISCDDETVNCYLID
 10 NNGFILVSEDYDTQTGDFEIEGAVMNKLLTMGSFKRITLYDYQAMCRANKESSDGAHGLL
 DPYNAFLSAVKWIMTELVLFLVEFNLCSWWHSDMTAKAQKLKQTLPCDTEYPAFVSERTI
 KETTGNIACEDCSKSFVIQQIPSSNLFMVVDSSCLCESVAPITMAPIEIRYNESLKCERL
 K

15 **12 - human amino acid sequence**

MAGPGSPRRASRGASALLAAALLYAALGDVVRSEQQIPLSVKLWASAFGGEIKSIAAKYS
 GSQLLQKKYKEYEKDVAIEEIDGLQLVKKLAKNMEEMFHKKSEAVRRLVEAAEEAHLKHEF
 DADLQYEFNAVLINERDKDGNFLELGKEFILAPNDHFNNLPVNISLSQVQVPTNMYNKDP
 AIVNGVYWSESLNKVFVDNFDPSLIWQYFGSAKGFFRQYPGIKWEPEENGVIAPDCRNR
 20 KWIYQAATSPKDVVILVDVSGSMKGLRLTIKQTVSSILDTLGDDDFNIIAYNEELHYVE
 PCLNGTLVQADRTNKEHFREHLDKLFAKGIGMLDIALNEAFNILSDFNHTGQGSICSQAIM
 LITDGAVDTYDTIFAKYNWPDRKVRIFTYILIGREAAAFADNLKWMACANKGFFTQISTLADV
 QENVMEYLHVLSRPKVIDQEHDVVWTEAYIDSTLTDDQGPVLMTTVAMPVFSKQNETRSKG
 ILLGVVGTDPVPKELLKTIPKYKLGIGHYAFAITNNGYILTHPELRLLYEKGKRRKPNYS
 25 SVDLSEVEWEDRDDVLRNAMVNRKTGKFSMEVKKTVDKGKRVLMVTNDYYYTDIKGTPFSL
 GVALSRGHGKYFFRGNVTIEEGLHDLEHPDVSLADEWSYCNTDLHPEHRHLSQLEAIKLYL
 KGKEPLLQCDKELIQEVLFDAVVSAPIEAYWTSALNKSSENSDKGVEVAFLGTRTGLSRIN
 LFVGAEQLTNQDFLKAGDKENIFNADHFPLWYRRAAEQIPGSFVYSIPFSTGPVNKSNVVT
 ASTSIQLLDERKSPVVAAVGIQMKLEFFQRKFWTASRQCASLDGKCSISCDDETVNCYLID
 30 NNGFILVSEDYDTQTGDFEIEGAVMNKLLTMGSFKRITLYDYQAMCRANKESSDGAHGLL
 DPYNAFLSAVKWIMTELVLFLVEFNLCSWWHSDMTAKAQKLKQTLPCDTEYPAFVSERTI
 KETTGNIACEDCSKSFVIQQIPSSNLFMVVDSSCLCESVAPITMAPIEIRYNESLKCERL
 KAKKIRRRRPESCHGFHPEENARECGGAP

35 **13 - human nucleotide sequence**

AGTGGCCTCCTGAGAAGCAGCTTGTTCTGTTGGGCTCCGAGAAGGTCTCCGACAGGAAGTTCC
 TGACACCTGAGGACGAGGCCAGCGTGTTACCCCTGGACCGCTTCCCGCTGTGGTACCGCCA
 GGCCTCAGAGCATCCTGCTGGCAGCTTCGTCTTCAACCTCCGCTGGGCAGAAGGACCAGAA

bioRxiv preprint doi: <https://doi.org/10.1101/055265>; this version posted September 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

AGTGC GGGTGAACCCATGGTGGTGACGGCAAGCACAGCTGTGGCGGTGACCGTGGACAAGA
 GGACAGCCATTGCTGCAGCCGCGGGCGTCCAAATGAAGCTGGAATTCCTCCAGCGCAAATT
 CTGGGCGGCAACGCGGCAGTGCAGCACTGTGGATGGGCCGTGCACACAGAGCTGCGAGGAC
 AGTGATCTGGACTGCTTCGTCATCGACAACAACGGGTTTCATTCTGATCTCCAAGAGGTCCC
 5 GAGAGACGGGAAGATTTCTGGGGGAGGTGGATGGTGCTGTCTGACCCAGCTGCTCAGCAT
 GGGGGTGTTCAGCCAAGTGACTATGTATGACTATCAGGCCATGTGCAAACCTCGAGTCAC
 CACCACAGTGCAGCCCAGCCCCTGGTTCAGCCCAATTTCTGCCTTCTTGACGGCGACCAGGT
 GGCTGCTGCAGGAGCTGGTGCTGTTCTGCTGGAGTGGAGTGTCTGGGGCTCCTGGTACGA
 CAGAGGGGCGGAGGCCAAAAGTGTCTTCCATCACTCCCACAAACACAAGAAGCAGGACCCG
 10 CTGCAGCCCTGCGACACGGAGTACCCCGTGTTCGTGTACCAGCCGGCCATCCGGGAGGCCA
 ACGGGATCGTGGAGTGCGGGCCCTGCCAGAAGGTATTTGTGGTGCAGCAGATTCCCAACAG
 TAACCTCCTCCTCCTGGTGACAGACCCACCTGTGACTGCAGCATCTTCCCACCAAGTG

14 - human nucleotide sequence

AGTGGCCTCCTGAGAAGCAGCTTGTTTCGTGGGCTCCGAGAAGGTCTCCGACAGGAAGTTCC
 TGACACCTGAGGACGAGGCCAGCGTGTTACCCCTGGACCGCTTCCCGCTGTGGTACCGCCA
 GGCCTCAGAGCATCCTGCTGGCAGCTTCGTCTTCAACCTCCGCTGGGCAGAAGGACCAGAA
 AGTGC GGGTGAACCCATGGTGGTGACGGCAAGCACAGCTGTGGCGGTGACCGTGGACAAGA
 GGACAGCCATTGCTGCAGCCGCGGGCGTCCAAATGAAGCTGGAATTCCTCCAGCGCAAATT
 20 CTGGGCGGCAACGCGGCAGTGCAGCACTGTGGATGGGCCGTGCACACAGAGCTGCGAGGAC
 AGTGATCTGGACTGCTTCGTCATCGACAACAACGGGTTTCATTCTGATCTCCAAGAGGTCCC
 GAGAGACGGGAAGATTTCTGGGGGAGGTGGATGGTGCTGTCTGACCCAGCTGCTCAGCAT
 GGGGGTGTTCAGCCAAGTGACTATGTATGACTATCAGGCCATGTGCAAACCTCGAGTCAC
 CACCACAGTGCAGCCCAGCCCCTGGTTCAGCCCAATTTCTGCCTTCTTGACGGCGACCAGGT
 25 GGCTGCTGCAGGAGCTGGTGCTGTTCTGCTGGAGTGGAGTGTCTGGGGCTCCTGGTACGA
 CAGAGGGGCGGAGGCCAAAAGTGTCTTCCATCACTCCCACAAACACAAGAAGCAGGACCCG
 CTGCAGCCCTGCGACACGGAGTACCCCGTGTTCGTGTACCAGCCGGCCATCCGGGAGGCCA
 ACGGGATCGTGGAGTGCGGGCCCTGCCAGAAGGTATTTGTGGTGCAGCAGATTCCCAACAG
 TAACCTCCTCCTCCTGGTGACAGACCCACCTGTGACTGCAGCATCTTCCCACCAAGTGCTG
 30 CAGGAGGCGACAGAAGTCAAATATAATGCCTCTGTCAAATGTGACCGGATGCGC

15 - human nucleotide sequence

AGTGGCCTCCTGAGAAGCAGCTTGTTTCGTGGGCTCCGAGAAGGTCTCCGACAGGAAGTTCC
 TGACACCTGAGGACGAGGCCAGCGTGTTACCCCTGGACCGCTTCCCGCTGTGGTACCGCCA
 35 GGCCTCAGAGCATCCTGCTGGCAGCTTCGTCTTCAACCTCCGCTGGGCAGAAGGACCAGAA
 AGTGC GGGTGAACCCATGGTGGTGACGGCAAGCACAGCTGTGGCGGTGACCGTGGACAAGA
 GGACAGCCATTGCTGCAGCCGCGGGCGTCCAAATGAAGCTGGAATTCCTCCAGCGCAAATT
 CTGGGCGGCAACGCGGCAGTGCAGCACTGTGGATGGGCCGTGCACACAGAGCTGCGAGGAC

09397550:091690

18 - human amino acid sequence

30 SGLLRSSLFVGSEKVS DRKFLTPED EASVFTLDRFPLWYRQASEHPAGSFVFNLRWAE GPE
SAGEPMVVTASTAVAVTVDKRTAIAAAAGVQM KLEFLQRKFWAATRQCSTVDGPCTQSCED
SDLDCFVIDNNGFILISKRSRETGRFLGEVDGAVLTQLLSMGVFSQVTMYDYQAMCKPSSH
HHSAAQPLVSPISAFLTATR WLLQELVLFLEWSVWGSWYDRGAEAKSVFHHSHKHKKQDP
LQPCDTEYPVFVYQPAIREANGIVECGPCQKV FVVQQIPNSNLLLLVTDPTCDCSI FPPVL
35 QEATEVKYNASVKCDRMR SQKLRRRPDSCHAFHPEENAODCGGAS

[illegible]

CGGGCAGCAGCCGCGCCGCGCAGGCGCTGCGGCCGTGCAGCCCGGAGGCCCTCGGGGAGA
AGGCGGCGGCGGAGGAGAGGCCGAGTTACCGCCCGCCGCCCGCGCCCCCAACCCGCGCG
CCGCCCGCGCCGCCGCGCCACTGCCCCCCTCCCCGCGGCGCCGCATCTTGAATGGAACATG
5 GCGGTGCCGGCTCGGACCTGCGGCGCCTCTCGGCCCGGCCAGCGCGGACTGCGCGCCCCCT
GGCCCCGGCTGCGGCCCCCACCCTGGCCCCGGCACCCGGCGCCCGACGTCCGGGCCCCCGCG
CCCCTGTGGCTGCTGCTGCCGCTTCTACCGCTGCTCGCCGCCCGCGCCTCTGCCTAC
AGCTTCCCCCAGCAGCACACGATGCAGCACTGGGCCCCGGCGTCTGGAGCAGGAGGTGACG
GCGTGATGCGGATTTTTGGAGGCGTCCAGCAGCTCCGTGAGATTTACAAGGACAACCGGAA
10 CCTGTTTCGAGGTACAGGAGAATGAGCCTCAGAAGTTGGTGGAGAAGGTGGCAGGGGACATT
GAGAGCCTTCTGGACAGGAAGGTGCAGGCCCTGAAGAGACTGGCTGATGCTGCAGAGAACT
TCCAGAAAGCACACCGCTGGCAGGACAACATCAAGGAGGAAGACATCGTGTA CTATGACGC
CAAGGCTGACGCTGAGCTGGACGACCCTGAGAGTGAGGATGTGGAAAGGGGGTCTAAGGCC
AGCACCTAAGGCTGGACTTCATCGAGGACCCAACTTCAAGAACAAGGTCAACTATTTCAT
15 ACGCGGCTGTACAGATCCCTACGGACATCTACAAAGGCTCCACTGTCATCCTCAATGAGCT
CAACTGGACAGAGGCCCTGGAGAATGTGTTTCATGGAAAACCGCAGACAAGACCCCACTG
CTGTGGCAGGTCTTCGGCAGCGCCACAGGAGTCACTCGCTACTACCCGGCCACCCCGTGGC
GAGCCCCCAAGAAGATCGACCTGTACGATGTCCGAAGGAGACCCTGGTATATCCAGGGGGC
CTCGTCACCCAAAGACATGGTCATCATCGTGGATGTGAGTGCGAGTGTGAGCGGCCTGACC
20 CTGAAGCTGATGAAGACATCTGTCTGCGAGATGCTGGACACGCTGTCTGATGATGACTATG
TGAATGTGGCCTCGTTCAACGAGAAGGCACAGCCTGTGTTCATGCTTACACACCTGGTGCA
GGCCAATGTGCGCAACAAGAAGGTGTTCAAGGAAGCTGTGCAGGGCATGGTGGCCAAGGGC
ACCACAGGCTACAAGGCCGGCTTTGAGTATGCCTTTGACCAGCTGCAGAACTCCAACATCA
CTCGGGCCAAC TGCAACAAGATGATCATGATGTTACGGATGGTGGTGAGGACCGCGTGCA
25 GGACGTCTTTGAGAAGTACAATTGGCCAAACCGGACGGTGCGCGTGTCTTACTTTCTCCGTG
GGGCAGCATAACTATGACGTCACACCGCTGCAGTGGATGGCCTGTGCCAACAAGGCTACT
ATTTTGAGATCCCTTCCATCGGAGCCATCCGCATCAACACACAGGAATATCTAGATGTGTT
GGGCAGGCCCATGGTGCTGGCAGGCAAGGAGGCCAAGCAGGTTTCAGTGGACCAACGTGTAT
GAGGATGCACTGGGACTGGGGTTGGTGGTAACAGGGACCCTCCCTGTTTTCAACCTGACAC
30 AGGATGGCCCTGGGGAAAAGAAGAACAGCTGATCCTGGGCGTGATGGGCATTGACGTGGC
TCTGAATGACATCAAGAGGCTGACCCCCAACTACACGCTTGGAGCCAACGGCTATGTGTTT
GCCATTGACCTGAACGGCTACGTGTTGCTGCACCCCAATCTCAAGCCCCAGACCACCAACT
TCCGGGAGCCTGTGACTCTGGACTTCCTGGATGCGGAGCTAGAGGATGAGAACAAGGAAGA
GATCCGTGCGGAGCATGATTGATGGCAACAAGGGCCACAAGCAGATCAGAACGTTGGTCAAG
35 TCCCTGGATGAGAGGTACATAGATGAGGTGACACGGAATAACCTGGGTGCCCTATAAGGA
GCACTAACTACAGCCTGGGGCTGGTGCTCCCACCCTACAGCACCTTCTACCTCCAAGCCAA
TCTCAGTGACCAGATCCTGCAGGTCAAGTATTTTGAGTTCTCTGCCAGCAGCTTTGAG
TCTGAAGGACACGTTTTTCATTGCTCCAGAGAGTACTGCAAGGACCTGAATGCCTCAGACA

ACAACACCGAGTTCCTGAAAACTTTATTGAGCTCATGGAGAAAGTGA CTCCAGACTCCAA
 GCAGTGCAACAACCTTCCTTCTGCACAACCTGATCTTGGACACGGGCATCACGCAGCAGCTG
 GTAGAGCGTGTGTGGAGGGACCAGGATCTCAACACGTACAGCCTACTGGCCGTGTTGCTG
 CCACAGACGGTGGCATCACCCGAGTCTTCCCCAACAAGGCAGCTGAGGACTGGACAGAGAA
 5 CCCTGAGCCCTTCAATGCCAGCTTCTACCGCCGAGCCTGGATAACCACGGTTATGTCTTC
 AAGCCCCCACACCAGGATGCCCTGTTAAGGCCGCTGGAGCTGGAGAATGACACTGTGGGCA
 TCCTCGTCAGCACAGCTGTGGAGCTCAGCCTAGGCAGGCGCACACTGAGGCCAGCAGTGGT
 GGGCGTCAAGCTGGACCTAGAGGCTTGGGCTGAGAAGTTCAAGGTGCTAGCCAGCAACCGT
 ACCCACCAAGACCAGCCTCAGAAGTGC GGCCCCAACAGCCACTGTGAGATGGACTGCGAGG
 10 TTAACAATGAGGACTTACTCTGTGTCTCATTTGATGATGGAGGATTCTCTGGTGTGTCAAA
 CCAGAACCATCAGTGGGACCAGGTGGGCAGGTTCTTCAGTGAGGTGGATGCCAACCTGATG
 CTGGCACTCTACAATAACTCCTTCTACACCCGCAAGGAGTCTTATGACTATCAGGCAGCCT
 GTGCCCCCTCAGCCCCCTGGCAACCTGGGTGCTGCACCCCGGGGTGTCTTTGTGCCCCACCGT
 TGCAGATTTCTTAACCTGGCCTGGTGGACCTCTGCTGCCGCTGGTCCCTGTTCCAGCAG
 15 CTTCTCTACGGCCTCATCTACCACAGCTGGTTCCAAGCAGACCCCGCGGAGGCCGAGGGGA
 GCCCCGAGACGCGGAGAGCAGCTGCGTCATGAAACAGACCCAGTACTACTTCGGCTCGGT
 AAACGCCTCCTACAACGCCATCATCGACTGCGGAACTGCTCCAGGCTGTTCCACGCGCAG
 AGACTGACCAACACCAATCTTCTCTTTGTGGTGGCCGAGAAGCCGCTGTGCAGCCAGTGCG
 AGGCTGGCCGGCTGCTGCAGAAGGAGACGCACTGCCAGCGGACGGCCCGGAGCAGTGTGA
 20 GCTAGTGCAGAGACCGCGATACCGGAGAGGCCCGCACATCTGCTTCGACTACAACGCGACA
 GAAGATACCTCAGACTGTGGCCGCGGGGCCCTCCTTCCCGCCGTGCTGGGCGTCTCTGGTCT
 CCCTGCAACTGCTGCTCCTCTGGGCCTGCCGCCCCGGCCGAGCCTCAAGTCTCTGTTCCA
 CGCCTCTCGCCGCTCTGAGCACCTTGCCCCACCCACCTCCACTCCCACCTCACCCGGCC
 TCTTCGCCTTTCCACCCCTCCTGCCCCACACTCCCCGCCTTAGAGCCTCGTCCCTCCCTCA
 25 CTGAAGGACCTGAGCTGGCCAGGCCCTGAGAGTCTGGTCTGCGCCTTGGGATGGGGAGTCC
 CAAAGCGGGACGCCGAGGTGTTTGGCACCCAAATCACATCTCACCTCCGAACCTGTTCAAG
 TGTCCCCAGACCCTTCTTGCTGCTGGGCTCCCCCAGTGGGATGGGACAGGGAGGCCACA
 CGCACTGGTGCCAAAACCAGGCCTCTGCTGCCGCCCTTCTTGAGGCTGCCTATGTTGGGG
 GGGACCTGCCTCAGCTGACCCGGCCTCTCTGCCCCACCCAAGCCCAAACCTGGTTTCTGT
 30 GAGAATAGTGGAGGAAGGTGAGATGGCCAGTTTGAAGCCTGTGCCTCCCAGCTTAAATCCT
 AGCAGGAGAGAGGCTCTGGGGCAGCCCCCATGGGCTCCTGCCCCCTTTCAGGCCTACAGCCA
 CATCCCCAAGCCCACCAGGTGTCAGGATAGTCACAGTGATACCAGTTCAGACACTACCCCA
 TATACACCTGGAACATTGAGGATGGAACTGGACTCACATTCGACATACCCCACTGGGCAC
 ACGCACAAACACACACACTATGGGGTGGGGTGGGTGTAGGGGCTTACAAAGCCTTACACAG
 35 GCGGAGGGGTGGTGGGAGGGTTGGCACCTGCACACTCCATCTCCTGCTCACCACCTGCCT
 CTAATCTGAGCTGCAGCCTGGCTGGTCTCTCCATTTCTAAAGCTGAATGTCAAACAGTGCC
 AAATGCTGGGGCAGGGGGTGAAGAACCCTCTGTCCACCCCTAGCCACCAGTGTCTCTCAA
 GTGCCCCCTCACCTCTCCAGGTGCTCATTGTAACCATTCTCTACTAGTGTGAGCCCCCAG

009760 0592680

TGGGACCACATGCCACTGCCTGCACCTTTCGGCAGAGGAACCCCCACCAGACATCACCCCTT
 TGCCTTAGCAGGGGTGACTTTGTCTCTCCTGGCTGGGCCATCCTTCCGCCAATCTGGCCCT
 TACACACTCAGGCCTGTGCCACTCCCTATCTCCTTCCCACCCCTACACACACACTCCCTG
 CTTGCAGGAGGCCAACTGTCCCTCCCTTGCTGAACACACACACACACACACACACAGG
 5 TGGGGACTGGGCACAGCTCTTCACACCATTCAATTCTGGTCATTTCCCCCAAAGGCATCCCA
 GCCTGGGGGCCAGTGGGGAAGTGGGGCAAGGGGATATAGTGATGGGGCTCAGATGGACTG
 GGAGGAGGGGGAGGGTGTATGATTAATTAATGGCTTCGTTAATTAATGTATGTTGCTTGT
 CGCTTTCTCAGTGTGTGTGTGTGGTCCATGCCCACTGCTGGTGCCAGGGTGGGTGTCCATG
 TGCACCCGGCTGGATGCCAGCTGTGTCTTCGGGGCGTGCGTGTAAGTGTAGTGTAGTC
 10 AGGTGCTCAATGGAGAATATAAACATATACAGAAAAATATATATTTTAAAGTTTAAAAAACA
 GAAAAACAGACAAAACAATCCCCATCAGGTAGCTGTCTAACCCCCAGCTGGGTCTAATCCT
 TCTCATTACCCACCCGACCTGGCTGCCCCCTCACCTTGGGCTGGGGGACTGGGGGGCCATTT
 CCTTTTCTCTGCCCTTTTTTTGTTGTTCTATTTTGTACAGACAAGTTGGAAAAACAACAGC
 GACAAAAAAGTCAAGAACTTTGTAAAATATCGTGTGTGTGATTCTTGTAATAATTTTTTC
 15 AAATGGTTTATTACAGAAGATCAGTTATTAAATAATGTTTCATATTTTCACTTC

20- human alpha2 delta-2 amino acid sequence

MAVPARTCGASRPGPARTARPWPGCGPHPGPTRRPTSGPPRPLWLLLPLLPLLAAPGASA
 YSFPQQHTMQHWARRLEQEVDGVMRIFGGVQQLREIYKDNRLFEVQENEPQKLVEKVAGD
 20 IESLLDRKVQALKRLADAAENFQKAHRWQDNIKEEDIVYYDAKADAELDDPESEDVERGSK
 ASTLRLDFIEDPNFKNKVNYSYAAVQIPTDIYKGSTVILNELNWTEALENVFMENRRQDPT
 LLWQVFGSATGVTRYYPATPWRAPKKIDLYDVRRRPWYIQGASSPKDMVIIVDVSGSVSGL
 TLKLMKTSVCEMLDTLSDDDYVNVASFNEKAQPVSCFTHLVQANVRNKKVFKEAVQGMVAK
 GTTGYKAGFEYAFDQLQNSNITRANCNKMIMMFTDGGEDRVQDVFEKYNWPNRTVRVFTFS
 25 VGQHNYDVTPLQWMACANKGYFFEIPSIGAIRINTQEYLDVLGRPMVLAGEAKQVQWNTV
 YEDALGLGLVVTGTLPVFNLTQDGPGEKKNLILGVMGIDVALNDIKRLTPNYTLGANGYV
 FAIDLNGYVLLHPNLKPQTTNFREPVTLDFLDAELEDENKEEIRRSIDGNKGHKQIRTLV
 KSLDERYIDEVTRNYTWVPIRSTNYSGLVLPPYSTFYLVQANLSDQILQVKYFEFLPSSF
 ESEGHVFIAPREYCKDLNASDNNTEFLKNFIELMEKVTPDSKQCNNFLLHNLILDTGITQQ
 30 LVERVWRDQDLNTYSLLAVFAATDGGITRVFPNKAEDWTENPEPFNASFYRRSLDNHGYV
 FKPPHQDALLRPLELENDTVGILVSTAVELSLGRRTLRPAVVGKLDLEAWAEKFKVLASN
 RTHQDQPPQKCGPNSHCCEMDCEVNNEDLLCVLIDDGGFLVLSNQHQQWDQVGRFFSEVDANL
 MLALYNNSFYTRKESYDYQAACAPQPPGNLGAAPRGVFVPTVADFLNLAWWTSAAAWSLFQ
 QLLYGLIYHSWFQADPAEAEAGSPETRESSCVMKQTQYYFGSVNASYNAIIDCGNCSRLFHA
 35 QRLTNTNLLFVVAEKPLCSQCEAGRLLQKETHCPADGPEQCELVQRPYRRGPHICFDYNA
 TEDTSDCGRGASFPPSLGVLVSLQLLLLLGLPPRPQPQLVHASRRL

0937550-091650

21 - human alpha2 delta-3 nucleotide sequence

TACTATAGGGCGGCCGCGAATTCGGCACGAGGCGGCGGAGCGGAGCAGGCAGCCCCGCG
 CGCTCGCCACCGCCCGCTCCGCGCAGCTCCCCGCGGCCGCTCTCGTCGCCGCCGCGAGCGG
 GCGCGTCGGAGGGAGCCCAGCATGGCCGGGCGGGCTCGCCGCGCCGCGCGTCCCGGGGGG
 5 CCTCGGCGCTTCTCGTGCCGCGCTTCTCTACGCCGCGCTGGGGGACGTGGTGCGCTCGGA
 GCAGCAGATACCGCTCTCCGTGGTGAAGCTCTGGGCCTCGGCTTTTGGTGGGGAGATAAAA
 TCCATTGCTGCTAAGTACTCCGGTTCAGCTTCTGCAAAAGAAATACAAAGAGTATGAGA
 AAGACGTTGCCATAGAAGAAATTGATGGCCTCCAAGTGGTAAAGAAGCTGGCAAAGAACAT
 GGAAGAGATGTTTACAAGAAGTCTGAGGCCGTGAGGCGTCTGGTGGAGGCTGCAGAAGAA
 10 GCACACCTGAAACATGAATTTGATGCAGACTTACAGTATGAATACTTCAATGCTGTGCTGA
 TAAATGAAAGGGACAAAGACGGGAATTTTTTGGAGCTGGGAAAGGAATTCATCTTAGCCCC
 AAATGACCATTTTAATAATTTGCCTGTGAACATCAGTCTAAGTGACGTCCAAGTACCAACG
 AACATGTACAACAAAGACCCTGCAATTGTCAATGGGGTTTATTGGTCTGAATCTCTAAACA
 AAGTTTTTGTAGATAACTTTGACCGTGACCCATCTCTCATATGGCAGTACTTTGGAAGTGC
 15 AAAGGGCTTTTTTAGGCAGTATCCGGGGATTAAATGGGAACCAGATGAGAATGGAGTCATT
 GCCTTCGACTGCAGGAACCGAAAATGGTACATCCAGGCAGCAACTTCTCCGAAAGACGTGG
 TCATTTTAGTTGACGTCAGTGGCAGCATGAAAGGACTCCGTCTGACTATCGCGAAGCAAAC
 AGTCTCATCCATTTTGGATACACTTGGGGATGATGACTTCTTCAACATAATTGCTTATAAT
 GAGGAGCTTCACTATGTGGAACCTTGCCTGAATGGAACCTTTGGTGCAAGCCGACAGGACAA
 20 ACAAAGAGCACTTCAGGGAGCATCTGGACAAACTTTTCGCCAAAGGAATTGGAATGTTGGA
 TATAGCTCTGAATGAGGCCCTTCAACATTCTGAGTGATTTCAACCACACGGGACAAGGAAGT
 ATCTGCAGTCAGGCCATCATGCTCATAACTGATGGGGCGGTGGACACCTATGATACAATCT
 TTGCAAAATACAATTGGCCAGATCGAAAGGTTTCGCATCTTACATACCTCATTGGACGAGA
 GGCTGCGTTTTGCAGACAATCTAAAGTGGATGGCCTGTGCCAACAAAGGATTTTTTACCCAG
 25 ATCTCCACCTTGGCTGATGTGCAGGAGAATGTCATGGAATACCTTCACGTGCTTAGCCGGC
 CCAAAGTCATCGACCAGGAGCATGATGTGGTGTGGACCGAAGCTTACATTGACAGCACTCT
 GACTGATGATCAGGGCCCCGTCCTGATGACCACTGTAGCCATGCCTGTGTTTAGTAAGCAG
 AACGAAACCAGATCGAAGGGCATTCTTCTGGGAGTGGTTGGCACAGATGTCCCAGTGAAAG
 AACTTCTGAAGACCATCCCCAAATACAAGTTAGGGATTACGGTTATGCCTTTGCAATCAC
 30 AAATAATGGRTATATCCTGACGCATCCGGAACCTCAGGCTGCTGTACGAAGAAGGAAAAAAG
 CGAAGGAAACCTAACTATAGTAGCGTTGACCTCTCTGAGGTGGAGTGGGAAGACCGAGATG
 ACGTGTTGAGAAATGCTATGGTGAATCGAAAGACGGGGAAGTTTTCCATGGAGGTGAAGAA
 GACAGTGGACAAAGGGAAACGGGTTTTGGTGTGACAAATGACTACTATTATACAGACATC
 AAGGGTACTCCTTTCAGTTTAGGTGTGGCGCTTTCAGAGGTCATGGGAAATATTTCTTCC
 35 GAGGGAATGTAACCATCGAAGAAGGCCTGCATGACTTAGAACATCCCGATGTGTCTTGGC
 AGATGAATGGTCTACTGCAACACTGACCTACACCCTGAGCACC GCCATCTGTCTCAGTTA
 GAAGCGATTAAGCTCTACCTAAAAGGCAAAGAACCTCTGCTCCAGTGTGATAAAGAATTGA
 TCCAAGAAGTCCTTTTTTGACGCGGTGGTGAGTGCCCCCATTTGAAGCGTATTGGACCAGCCT

093750:0526E60

GGCCCTCAACAAATCTGAAAATTCTGACAAGGGCGTGGAGGTTGCCTTCCTCGGCACTCGC
 ACGGGCCTCTCCAGAATCAACCTGTTTGTCTGGGGCTGAGCAGCTCACCAATCAGGACTTCC
 TGAAAGCTGGCGACAAGGAGAACATTTTAAACGCAGACCATTTCCCTCTCTGGTACCGAAG
 AGCCGCTGAGCAGATTCCAGGGAGCTTCGTCTACTCGATCCCATTGAGCACTGGACCACTC
 5 AATAAAAGCAATGTGGTGACAGCAAGTACATCCATCCAGCTCCTGGATGAACGGAAATCTC
 CTGTGGTGGCAGCTGTAGGCATTGAGATGAACTTGAATTTTTTCCAAAGGAAGTTCTGGAC
 TGCCAGCAGACAGTGTGCTTCCCTGGATGGCAAATGCTCCATCAGCTGTGATGATGAGACT
 GTGAATTGTTACCTCATAGACAATAATGGATTTATTTTGGTGTCTGAAGACTACACACAGA
 CTGGAGACTTTTTTGGTGAGATCGAGGGAGCTGTGATGAACAAATTGCTAACAATGGGCTC
 10 CTTTAAAAGAATTACCCTTTATGACTACCAAGCCATGTGTAGAGCCAACAAGGAAAGCAGC
 GATGGCGCCCATGGCCTCCTGGATCCTTATAATGCCTTCCTCTCTGCAGTAAATGGATCA
 TGACAGAACTTGTCTTGTTCCTGGTGGAATTTAACCTCTGCAGTTGGTGGCACTCCGATAT
 GACAGCTAAAGCCCAGAAATTGAAACAGACCCTGGAGCCTTGTGATACTGAATATCCAGCA
 TTCGTCTCTGAGCGCACCATCAAGGAGACTACAGGGAATATTGCTTGTGAAGACTGCTCCA
 15 AGTCCCTTGTGTCATCCAGCAAATCCCAAGCAGCAACCTGTTTCATGGTGGTGGTGGACAGCAG
 CTGCCCTCTGTGAATCTGTGGCCCCCATCACCATGGCACCATTGAAATCAGGTATAATGAA
 TCCCTTAAGTGTGAACGTCTAAAGGCCAGAAAGATCAGAAGGCGCCAGAAATCTTGTGATG
 GCTTCCATCCTGAGGAGAATGCAAGGGAGTGTGGGGGTGCGCCGAGTCTCCAAGCCCAGAC
 AGTCTCCTTCTGCTCCCTCTGCTTTTGTGCTCTTCTCAAGGTGACACTGACTGAGATGT
 20 TCTCTTACTGACTGAGATGTTCTCTTGGCATGCTAAATCATGGATAAACTGTGAACCAAAA
 TATGGTGCAACATACGAGACATGAATATAGTCCAACCATCAGCATCTCATCATGATTTTAA
 ACTGTGCGTGATATAAACTCTTAAAGATATGTTGACAAAAAGTTATCTATCATCTTTTTTAC
 TTTGCCAGTCATGCAAATGTGAGTTTGCCACATGATAATCACCTTCATCAGAAATGGGAC
 CGCAAGTGGTAGGCAGTGTCCCTTCTGCTTGAAACCTATTGAAACCAATTTAAAACTGTGT
 25 ACTTTTTTAAATAAAGTATATTAAATCATAAAAAAAAAAAAAAAAAAAAA

22 - human alpha2 delta-3 amino acid sequence

MAGPGSPRRASRGASALLAAALLYAALGDVVRSEQQIPLSVVKLWASAFGGEIKSIAAKYS
 GSQLLQKKYKEYEKDVAIEEIDGLQLVKKLAKNMEEMFHKKSEAVRRLVEAAEEAHLKHEF
 30 DADLQYEFNAVLINERDKDGNFLELGKEFILAPNDHFNNLPVNISLSDVQVPTNMYNKDP
 AIVNGVYWSESLNKVFVDNFDRLPSLIWQYFGSAKGFFRQYPGIKWEPDENGVIADFCDNR
 KWIQAATSPKDVVILVDVSGSMKGLRLTIKQTVSSILDTLGDDDDFFNI IAYNEELHYVE
 PCLNGTLVQADRTNKEHFREHLDKLFAKGIGMLDIALNEAFNILSDFNHTGQGSICSQAIM
 LITDGAVDITYDTIFAKYNWPDRKVRIFTYLGREAAAFADNLKWMACANKGFFTQISTLADV
 35 QENVMEYLHVLSPKVIDQEHDDVVWTEAYIDSTLTDDQGPVLMTTVAMPVFSKQNETRSKG
 ILLGVVGTDPVKELLKTIPKYKLGIGHYAFAITNNGYILTHPELRLLYEEGKKRRKPNYS
 SVDLSEVEWEDRDDVLRNAMVNRKTGKFSMEVKKTVDKGKRVLMVTNDYYYTDIKGTPFSL
 GVALSRGHGKYFFRGNVITIEGLHDLHPDVS LADEWSYCNTDLHPEHRHLSQLEAIKLYL

KGKEPLLQCDKELIQEVLFDVVSAPIEAYWTSALNKSSENSDKGVEVAFLGTRTGLSRIN
 LFVGAEQLTNQDFLKAGDKENIFNADHFPLWYRRAAEQIPGSFVYSIPFSTGPNKSNVVT
 ASTSIQLLDERKSPVVAAGVIQMKLEFFQRKFWTASRQCASLDGKCSISCDDTVNCYLID
 NNGFILVSEDTQTGDFGEIEGAVMNKLLTMGSFKRITLYDYQAMCRANKESSDGAHGLL
 5 DPYNAFLSAVKWIMTELVLFLVEFNLCSSWWHSDMTAKAQKLKQTLPCDTEYPAFVSERTI
 KETTGNIACEDCSKSFVIQQIPSSNLFMVVDSSCLCESVAPITMAPIEIRYNESLK CERL
 KAQKIRRRPESCHGFHPEENARECGGAPSLQAQTVLLLLLPLLLMLFSR

23 - human amino acid sequence

10 MAVPARTCGASRPGPARTARPWPGCGPHPGTTRRPTSGPPRPLWLLLPLLLPLLAAPGASA
 YSFPQQHTMQHWARRLEQEVDGVMRIFGGVQQLREIYKDNRNLFQENEPQKLVEKVAGD
 IESLLDRKVQALKRLADAAENFQKAHRWQDNIKEEDIVYYDAKADAELDDPESEDVERGSK
 ASTLRDLFIEDPNFKNKVNSYAAVQIPTDIYKGSTVILNELNWTEALENVFMENRRQDPT
 LLWQVFGSATGVTRYYPATPWRAPKKIDLYDVRRRPWYIQGASSPKDMVIIVDVSGSVSGL
 15 TLKLMKTSVCEMLDTLSDDDYVNVAASFNEKAQPVSCFTHLVQANVRNKKVFKEAVQGMVAK
 GTTGYKAGFEYAFDQLQNSNITRANCNKMIMMFTDGGEDRVQDVFEKYNWPNRTVRVFTFS
 VGQHNYDVTPLQWMACANKGYIFEIPSIGAIRINTQEYLDVLGRPMVLAGEAKQVQWTVN
 YEDALGLGLVVTGTLPVFNLTQDGPGEKKQLILGVMGIDVALNDIKRLTPNYTLGANGYV
 FAIDLNGYVLLHPNLKPQTTNFREPVTLDFLDAELEDENKEEIRRSIDGNKGHKQIRTLV
 20 KSLDERYIDEVTRNYTWVPIRSTNYSGLVLPPYSTFYQLQANLSDQILQVKYFEFLPSSF
 ESEGHVFIAPREYCKDLNASDNNTEFLKNFIELMEKVTPDSKQCNNFLLHNLILDGTITQQ
 LVERVWRDQDLNTYSLLAVFAATDGGITRVFPNKAEDWTENPEPFNASFYRRSLDNHGIV
 FKPPHQDALLRPLELENDTVGILVSTAVELSLGRRTLRAVVGVKLDLEAWAEKFKVLASN
 RTHQDQPQKCGPNSHCEMDCEVNNEDLLCVLIDDGGFLVLSNQNHQWDQVGRFFSEVDANL
 25 MLALYNNSFYTRKESYDYQAACAPQPPGNLGAAPRGVFVPTVADFLNLAWWTSAAAWSLFQ
 QLLYGLIYHSWFQADPAEAGSPETRESSCVMKQTQYYFGSVNASYNAIIDCGNCSRLFHA
 QRLTNTNLLFVVAEKPLCSQCEAGRLLQKETHCPADGPEQCELVQRPRYRRGPHICFDYNA
 TEDTSDCGRGAHHHHHH

30 24- mouse amino acid sequence

MAGPGSLCCASRGASALLATALLYAALGDVVRSEQQIPLSVVKLWASAFGGEIKSIAAKYS
 GSQLLQKKYKEYEKDVAIEEIDGLQLVKKLAKIMEEMFHKKSEAVRRLVEAAEEAHLKHEF
 DADLQYEFYFNAVLINERDKDGNFLELGKEFILAPNDHFNNLPVNISLSDVQVPTNMYNKDP
 AIVNGVYWSESLNKVFVDNFDNRDPSLIWQYFGSAKGFFRQYPGIKWEPDENGVIADFCDNR
 35 KWIQAATSPKDVVILVDVSGSMKGLRLTIKQTVSSILDTLGDDDFNIITYNEELHYVE
 PCLNGTLVQADRTNKEHFRHLDKLFAKGIGMLDIALNEAFNILSDFNHTGQGSICSQAIM
 LITDGAVDTYDTIFAKYNWPDRKVRIFTYILIGREAAAFADNLKWMACANKGFFTQISTLADV
 QENVMEYLHVLSRPKVIDQEHDDVVWTEAYIDSTLPQAQKLADDQGLVLMTTVAMPVFSKQN

009750-001600

[illegible]

TCGCCACCATGGCGGTGCCGGCTC

TCGCCACCATGGCGGTGCCGGCTC

TCGGAATTCCTCAGTGATGGTGATGGTGATGGGCCCCGCGGCCACAGTC

TCGGAATTCCTCAGTGATGGTGATGGTGATGGGCCCCGCGGCCACAGTC

TCGCCACCATGGCCGGGCGGGGC

TCGCCACCATGGCCGGGCGGGGC

28 - nucleotide sequence
TCTCAGTGATGGTGATGGTGATGCGATGCACCCCCACACTCTC

TCTCAGTGATGGTGATGGTGATGCGATGCACCCCACACTCTC